

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2003, 05:23:15 ; Search time 963 Seconds
(without alignments)
10147.862 Million cell updates/sec

Title: US-09-974-973-1
Perfect score: 3474
Sequence: 1 gtactgctatcaccccttg.....tgatcgctgcgtttctctaa 3474

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3474	100.0	3474	6	AX453601	AX453601 Sequence
2	3474	100.0	3474	6	AX453603	AX453603 Sequence
3	3398.8	97.8	3621	6	AR123974	AR123974 Sequence
4	3398.8	97.8	3637	1	AF038548	AF038548 Corynebact
5	3398.8	97.8	3728	1	CGPYC	Y09548 Corynebacte
6	3398.8	97.8	3728	6	A97276	A97276 Sequence 1
7	3398.8	97.8	332050	1	AP005276	AP005276 Corynebac
8	3398.8	97.8	349980	6	AX127145	AX127145 Sequence
9	3394	97.7	3657	1	AF503915	AF503915 Corynebact
10	3344.8	96.3	3420	6	AX120849	AX120849 Sequence
11	2256.2	64.9	4851	1	AB083299	AB083299 Corynebact
12	1564.8	45.0	1719	6	AX064993	AX064993 Sequence
13	1351.8	38.9	1406	6	AX064995	AX064995 Sequence
14	1351	38.9	40104	1	SCF11	AL132662 Streptomy
15	1329.8	38.3	4218	1	AF262949	AF262949 Mycobacte
16	1283.2	36.9	17992	1	AE007125	AE007125 Mycobacte
17	1283.2	36.9	43523	1	MTCY349	283018 Mycobacteri
18	1238.4	35.6	31175	1	MTU000024	U000024 Mycobacteri
19	1238.4	35.6	31175	1	MSGY2	AD000009 Mycobacte
20	1020.2	29.4	1083	6	AX064989	AX064989 Sequence
21	1020.2	29.4	1083	6	AX064991	AX064991 Sequence
22	883	25.4	939	6	AX064985	AX064985 Sequence
23	883	25.4	939	6	AX064987	AX064987 Sequence
24	695.6	20.0	3916	8	AF097728	AF097728 Aspergill
25	691.6	19.9	3559	5	AF509529	AF509529 Gallus ga
26	688.6	19.8	3953	9	HSU04641	U04641 Human pyruv
27	688.6	19.8	4058	9	BC011617	BC011617 Homo sapi
28	687	19.8	3960	9	HSU0891	U30891 Human pyruv
29	684.2	19.7	4067	10	MUSMPYR	L09192 Mus musculu
30	682.6	19.6	4024	10	RNU36585	U36585 Rattus norv
31	681	19.6	3945	6	AX401890	AX401890 Sequence
32	681	19.6	3945	10	RNU32314	U32314 Rattus norv
33	674.2	19.4	4017	6	AX410383	AX410383 Sequence
34	674.2	19.4	4017	9	S72370	S72370 pyruvate ca
35	670.4	19.3	6513	8	AN19972	AJ009972 Aspergill
36	665.2	19.1	4371	3	AF132152	AF132152 Drosophil
37	648	18.7	42926	1	MLCB1243	AL023635 Mycobacte
38	648	18.7	348950	1	MLEPRTN7	AL583923 Mycobacte
39	647.8	18.6	38477	1	MSGB1723CS	L78825 Mycobacteri
40	640.4	18.4	3800	1	REU51439	U51439 Rhizobium e
41	640.4	18.4	349116	1	AP003003	AP003003 Mesorhizo
42	639	18.4	4000	5	AF295372	AF295372 Danio rer
43	622	17.9	11295	1	AE009469	AE009469 Brucella
44	611	17.6	3727	1	D83706	D83706 Bacillus st
45	609	17.5	7076	1	AE009219	AE009219 Agrobacte

ALIGNMENTS

RESULT 1
AX453601
LOCUS AX453601
DEFINITION Sequence 1 from Patent WO0231158.
ACCESSION AX453601
VERSION AX453601.1 GI:21712841
KEYWORDS
SOURCE Corynebacterium glutamicum.
ORGANISM Corynebacterium glutamicum
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Corynebacteriaceae;
Corynebacterium.
REFERENCE 1
AUTHORS Hanke, P.D.
TITLE Feedback-resistant pyruvate carboxylase gene from corynebacterium

JOURNAL	Patent: WO 0231158-A 1 18-APR-2002;
FEATURES	ARCHER-DANIELS-MIDLAND COMPANY (US)
source	Location/Qualifiers
	1..3474
	/organism="Corynebacterium glutamicum"
	/db_xref="taxon:1718"
CDS	1..3474
	/note="unnamed protein product"
	/transl_table=1
	/codon_start=1
	/protein_id="CAD38136.1"
	/db_xref="gi:21712842"
	/translation="MTATITGGLLKGIITLVSTSTLPFAFKLIVANRGEIAVRA FRALEGTAAVAVIPREDRSGFHRSEAFSEAVRIGTEGSPVKAYLIDLEIIGAIAKVKV ADALVPGGFLSENAOLARCAENGITTFIGTPEVLDTLGDKSRVATAAKKAGI.PVLA ESTSKNIDILVKSAGQTYPIFKVAVAGGGRGMRVSSPDLRKLKATASAREAA PGDSVVERAVINPOHIEVOILGDRFTGVVHLVERDCSLORRKHQKVEIAPAHLDP ELDRICADAVKFCRSIGYOGACTVFLEDEKGNHVFLEMPRILOVEHTVTEEVTD LVKQMLBLAGATLKEILGKIDOKIKTHGAALOCRIITREDPNNGFRPDTGITAVRSPG CAGVRLDCAQLCGEITAFHEDSMIVKMTGCGDFETAVARAORALAFETTSVATNIG FLRALREEDFTSKRIATGFIQSHPHLLKPPADDEQSRIDYLDVTVAKPKGVREK DYAPIDKLPNIKDLPIPRGSRDLKQLGPAAPARDLREQDALATDTTFRDAHQSLIL ATVRSFALKPAEAAVAKLTPELLSVEAMGGATYDVAMRFEDFPWKLDLELRAMEN VNIQMLRGNTVGTTPDSVCRKAFVKAASGVDDIPRIFDALNDVDSOMKPAIDAVIL ENTNAEVAWAYSGDLSDENEKLYTLDITLKMAEEIVKSGAHILAIKIMAGLLRPPAA VTKLVLRREFDLPVHVHTHTAGGQLATYFAAAGADAVDGNASPLSGTSQPSEL SAIVAAATLGLLEAVSDLEPYWEAVRGLXLDLPFESGTPGPTGKRVKHEIPGGQ LSNRQAATLGLADRFELIEDNYAVNEMLRPTKVTPSSKVGDLALHLVAGVDDP ADFAADPKQYDIPDSVIAFLRGLGNP PGMPPEPLRTALRGSESGKAPLTVPEEQ AHLDDADS KERNLSRLFLPKPTEFLEHRRFNGTNSALDDREFFYGLVEGRETLIR LPDVRTLLRLDLAISPPDDKGMRYNVANVNGQIRPMRVDRSVESVTATAEKADSSN KGHAAPFVGVVTVVTAEGDEVKAGDAVAIIEMKMEATITASVDGKIERYVVPATK VEGDLIVVGS"
BASE COUNT	749 a 1013 c 950 g 762 t
ORIGIN	
Query Match	100.0%; Score 3474; DB 6; Length 3474;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 3474; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 GTCACTGCTATCACCCCTGGCGGTCTCTTTGTTGAAAGGAATAATTAATCTAGTGTGCTGACT 60
Db	1 GTCACTGCTATCACCCCTGGCGGTCTCTTTGTTGAAAGGAATAATTAATCTAGTGTGCTGACT 60
Qy	61 CACACATCTCAACGCTTCCAGCATTCAAAAGATCTTGTAGCAACCGCGGGAATC 120
Db	61 CACACATCTCAACGCTTCCAGCATTCAAAAGATCTTGTAGCAACCGCGGGAATC 120
Qy	121 GCGGTGCGTCTTCCGTGACACTCGAAACCGGTGCGACGCGAGCTAGTATTACCC 180
Db	121 GCGGTGCGTCTTCCGTGACACTCGAAACCGGTGCGACGCGAGCTAGTATTACCC 180
Qy	181 CGTGAAGATCGGGGATCAFTCCACCGCTCTTTGCTTCTGAAGCTGTCGCAATGCTACT 240
Db	181 CGTGAAGATCGGGGATCAFTCCACCGCTCTTTGCTTCTGAAGCTGTCGCAATGCTACT 240
Qy	241 GAAGGCTCAACGATCAAGCGTACCTGGACATCGATGAAATTTATCGGTGCGAGCTAAAAA 300
Db	241 GAAGGCTCAACGATCAAGCGTACCTGGACATCGATGAAATTTATCGGTGCGAGCTAAAAA 300
Qy	301 GTTAAAGCAGATGCTATTATACCGGGATATGGCTTCTGTTCTGAAATGCGCAGCTTGCC 360
Db	301 GTTAAAGCAGATGCTATTATACCGGGATATGGCTTCTGTTCTGAAATGCGCAGCTTGCC 360
Qy	361 CGCGAGTGCAGGAAACCGCATTAATTTTATTTGCCCCAACCCAGAGGTTCTTGATCTC 420
Db	361 CGCGAGTGCAGGAAACCGCATTAATTTTATTTGCCCCAACCCAGAGGTTCTTGATCTC 420
Qy	421 ACCGGTGATAGTCTCTGCGGTAAACCGCGCGAAGAGGCTGCTGCGCAGTTTGGCG 480
Db	421 ACCGGTGATAGTCTCTGCGGTAAACCGCGCGAAGAGGCTGCTGCGCAGTTTGGCG 480
Qy	481 GAATCCACCCCGAGCAAAAACATCGATGACATCGTTAAAGCGCTGAAGGCGCAGACTTAC 540
Db	GAATCCACCCCGAGCAAAAACATCGATGACATCGTTAAAGCGCTGAAGGCGCAGACTTAC 540
Qy	541 CCCATCTTTGTAAGGAGGAGTTCGCGGTGGTGGGACGCGGTATGCGCTTTGTTCTTCA 600
Db	CCCATCTTTGTAAGGAGGAGTTCGCGGTGGTGGGACGCGGTATGCGCTTTGTTCTTCA 600
Qy	601 CCTGATGAGCTCCGCAAAATTTGGCAACAGAAAGCATCTCGTGAAGCTGAAGCGGCAATTCGGC 660
Db	CCTGATGAGCTCCGCAAAATTTGGCAACAGAAAGCATCTCGTGAAGCTGAAGCGGCAATTCGGC 660
Qy	661 GACGGTTCGGTATATGTCGAAACGTCGTGTGATTAAACCCCGACACATTTGAAGTGCAGATC 720
Db	GACGGTTCGGTATATGTCGAAACGTCGTGTGATTAAACCCCGACACATTTGAAGTGCAGATC 720
Qy	721 CTTTGGCGATCCGACCTGAGAGGTTTGTACACCTTTATGAACGTCGCTGCTACTGCGAGGT 780
Db	CTTTGGCGATCCGACCTGAGAGGTTTGTACACCTTTATGAACGTCGCTGCTACTGCGAGGT 780
Qy	781 CGTCACCAAAAAGTTGTCGAAATTTGCCGACGACATTTGGATCCAGAACTGCGTGTAT 840
Db	CGTCACCAAAAAGTTGTCGAAATTTGCCGACGACATTTGGATCCAGAACTGCGTGTAT 840
Qy	841 CGCATTTCGCGGATGACGTAAGATTTCTGCCGCTCCATTGGTTACAGGGCGCGGAACC 900
Db	CGCATTTCGCGGATGACGTAAGATTTCTGCCGCTCCATTGGTTACAGGGCGCGGAACC 900
Qy	901 GTGGAATTTCTTGGTTCGATGAAAAAGGGCAACACGTTTTCATCGAAATGAAACCCAGCTATC 960
Db	GTGGAATTTCTTGGTTCGATGAAAAAGGGCAACACGTTTTCATCGAAATGAAACCCAGCTATC 960
Qy	961 CAGGTTGAGCAACACCGTGACTGAAAGAGTCAACCGAGGTGACCTGGTGAAGGCGCAGATG 1020
Db	CAGGTTGAGCAACACCGTGACTGAAAGAGTCAACCGAGGTGACCTGGTGAAGGCGCAGATG 1020
Qy	1021 CGCTTGGCTGCTGGTGCACACTTTGAAGGAATTTGGGTCTCGACCAAGATAAGATCAAGACC 1080
Db	CGCTTGGCTGCTGGTGCACACTTTGAAGGAATTTGGGTCTCGACCAAGATAAGATCAAGACC 1080
Qy	1081 CACGCTGACGACATGTCAGTCCGATCACCACGGAAGATCCAAAACAAAGGCTTCGCGCCA 1140
Db	CACGCTGACGACATGTCAGTCCGATCACCACGGAAGATCCAAAACAAAGGCTTCGCGCCA 1140
Qy	1141 GATACCGGAACATATCACCGGTACCGCTCACCGAGCGGAGCTGGCTTCTGTTGACGGT 1200
Db	GATACCGGAACATATCACCGGTACCGCTCACCGAGCGGAGCTGGCTTCTGTTGACGGT 1200
Qy	1201 GCAGCTCAGCTCGGTGGGGAATCACCGACACTTTGACTCCATCTGCTGAAATGACC 1260
Db	GCAGCTCAGCTCGGTGGGGAATCACCGACACTTTGACTCCATCTGCTGAAATGACC 1260
Qy	1261 TCCGCTGGTTCGGAATTTGAAACTGCTGTTGCTGCTGACAGCGCGCTTGGCTGAGTTC 1320
Db	TCCGCTGGTTCGGAATTTGAAACTGCTGTTGCTGCTGACAGCGCGCTTGGCTGAGTTC 1320
Qy	1321 ACCGTGCTCGGTGTCACCAACAAATTTGCTTGGTGGCTTGGTGGGGAAGAGGAC 1380
Db	ACCGTGTCTGGTGTGCAACCAACAAATTTGCTTGGTGGCTTGGTGGGGAAGAGGAC 1380
Qy	1381 TTCACTTCAAGCGCATCGCCACCGGATTTATCGCGATTCACCCACACCTCTCTCAGGCT 1440
Db	TTCACTTCAAGCGCATCGCCACCGGATTTATCGCGATTCACCCACACCTCTCTCAGGCT 1440
Qy	1441 CCACCTGCGGATGATGACGAGGACGATCCTTGGATTTACTTTGGCAGATGTACCCGTGAAC 1500
Db	CCACCTGCGGATGATGACGAGGACGATCCTTGGATTTACTTTGGCAGATGTACCCGTGAAC 1500
Qy	1501 AAGCTCTAGTGTGCGGTCCAAAAGGATTTTGCAGACCAATCGAATAGTGCACCAATC 1560
Db	AAGCTCTAGTGTGCGGTCCAAAAGGATTTTGCAGACCAATCGAATAGTGCACCAATC 1560
Qy	1561 AAGGATCTGCCACTGCCACGCGTTCCCGGTGACCGCTGAAGCAGCTTGGCCACCGCG 1620
Db	AAGGATCTGCCACTGCCACGCGTTCCCGGTGACCGCTGAAGCAGCTTGGCCACCGCG 1620

Db 1561 AAGNATCTGCCACTGCCAGCGGTTCCGTGACCGCCTGAAGCAGCTTGGCCCGAGCGG 1620
Qy 1621 TTGTGCTGTGATCTCCGTGACGAGCGCACTGCGAGTTACTGATACACACTTCCGCGAT 1680
Db 1621 TTGTGCTGTGATCTCCGTGACGAGCGCACTGCGAGTTACTGATACACACTTCCGCGAT 1680
Qy 1681 GCACACAGCTTTGCTTGGACCGGAGTCCGCTCATTCGCACTGAAGCCTGCGCGCAGAG 1740
Db 1681 GCACACAGCTTTGCTTGGACCGGAGTCCGCTCATTCGCACTGAAGCCTGCGCGCAGAG 1740
Qy 1741 GCCGTGCAAAAGCTGACTCTCTGAGCTTTGTCCTGTGAGGCTGCGGGCGGCGGACCTAC 1800
Db 1741 GCCGTGCAAAAGCTGACTCTCTGAGCTTTGTCCTGTGAGGCTGCGGGCGGCGGACCTAC 1800
Qy 1801 GATGTGGCGATGCGTTTCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGGAG 1860
Db 1801 GATGTGGCGATGCGTTTCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGGAG 1860
Qy 1861 GCGATGCCGATGTAAACATTTAGATGCTGCTTCCGCGCCGCAACACCGTGGGATACACC 1920
Db 1861 GCGATGCCGATGTAAACATTTAGATGCTGCTTCCGCGCCGCAACACCGTGGGATACACC 1920
Qy 1921 CCGTACCAGACTCCGCTGCGCGCGTTTGTAAAGGAAGCTGCCAGCTCCGCGGTGAC 1980
Db 1921 CCGTACCAGACTCCGCTGCGCGCGTTTGTAAAGGAAGCTGCCAGCTCCGCGGTGAC 1980
Qy 1981 ATCTTCCGCACTTTGAGCGCTTTAAGCAAGTCTCCAGATGCGTCCAGCAATCGACGCA 2040
Db 1981 ATCTTCCGCACTTTGAGCGCTTTAAGCAAGTCTCCAGATGCGTCCAGCAATCGACGCA 2040
Qy 2041 GTCTTGAGACCAACCGCGGTAGCCGAGTGGCTATGGCTTATTTGTTGATCTCTCT 2100
Db 2041 GTCTTGAGACCAACCGCGGTAGCCGAGTGGCTATGGCTTATTTGTTGATCTCTCT 2100
Qy 2101 GATCAAAATGAAAGCTCTACACCTGATTTACTACCTTAAAGATGGCAGAGAGATCGTC 2160
Db 2101 GATCAAAATGAAAGCTCTACACCTGATTTACTACCTTAAAGATGGCAGAGAGATCGTC 2160
Qy 2161 AAGTCTGCGCTCACTTTGCGCATTAAGGATATGGCTGCTGCTTCCGCCAGCTGCG 2220
Db 2161 AAGTCTGCGCTCACTTTGCGCATTAAGGATATGGCTGCTGCTTCCGCCAGCTGCG 2220
Qy 2221 GTACCAAGCTGTGTCACCGCACTGCGCGGTGAATTCGATCTGCCAGTGCACTGCACACC 2280
Db 2221 GTACCAAGCTGTGTCACCGCACTGCGCGGTGAATTCGATCTGCCAGTGCACTGCACACC 2280
Qy 2281 CACGACACTGCGGTGCGCAGTTCGCTACTTTCGTCAGCTCAAGCTGGTGCGAGAT 2340
Db 2281 CACGACACTGCGGTGCGCAGTTCGCTACTTTCGTCAGCTCAAGCTGGTGCGAGAT 2340
Qy 2341 GCTGTGACCGTGTCTCGCACCACTGTCTGGCACCACTCCAGCCATCCCTGCTGCCC 2400
Db 2341 GCTGTGACCGTGTCTCGCACCACTGTCTGGCACCACTCCAGCCATCCCTGCTGCCC 2400
Qy 2401 ATTGTGCTGATCTGCGCGCAACCGCTGCGATACCGGTTTGGCCTCGAGGCTGTTTCT 2460
Db 2401 ATTGTGCTGATCTGCGCGCAACCGCTGCGATACCGGTTTGGCCTCGAGGCTGTTTCT 2460
Qy 2461 GACCTCGAGCGCTACTGGGAAGCTGTCGCGGAGTGTACCTGCCATTTGAGTCTGGAACC 2520
Db 2461 GACCTCGAGCGCTACTGGGAAGCTGTCGCGGAGTGTACCTGCCATTTGAGTCTGGAACC 2520
Qy 2521 CCAGGCCCAACCGGTGCGCTTACCGCCACAGAAATCCAGCGGACAGTGTGTCACAACTG 2580
Db 2521 CCAGGCCCAACCGGTGCGCTTACCGCCACAGAAATCCAGCGGACAGTGTGTCACAACTG 2580
Qy 2581 CGTGACAGGCCACCGCACTGGGCTTGTGATCGCTTCGAGCTCATCGAAGACAACTAC 2640
Db 2581 CGTGACAGGCCACCGCACTGGGCTTGTGATCGCTTCGAGCTCATCGAAGACAACTAC 2640
Qy 2641 GCACCGGTTAATGATGCTGGGAGCGCCCAACCAAGGTACACCCATCTCCCAAGTTGTT 2700
Db 2641 GCACCGGTTAATGATGCTGGGAGCGCCCAACCAAGGTACACCCATCTCCCAAGTTGTT 2700

Qy 2701 GGGCACTCTCGCACTCCACCTGGTTGGTGGGGTGTAGATCCAGCAGACTTTTGTCTGCAGAC 2760
Db 2701 GGGCACTCTCGCACTCCACCTGGTTGGTGGGGTGTAGATCCAGCAGACTTTTGTCTGCAGAC 2760
Qy 2761 CCACAAAGTACGACATCCAGACTCTGTCTATCGCTTCTGCGCGCGAGCTTGGTAAAC 2820
Db 2761 CCACAAAGTACGACATCCAGACTCTGTCTATCGCTTCTGCGCGCGAGCTTGGTAAAC 2820
Qy 2821 CTTCCAGTGGCTGGCCAGAACCACTGCGCACCCGCGCACTTGAAGGCCGCTCCGAAGGC 2880
Db 2821 CTTCCAGTGGCTGGCCAGAACCACTGCGCACCCGCGCACTTGAAGGCCGCTCCGAAGGC 2880
Qy 2881 AAGCAGCTCTGACGGAAGTTCCTGAGGAAGAGCAGCGCACCTCGACGCTGATGATTC 2940
Db 2881 AAGCAGCTCTGACGGAAGTTCCTGAGGAAGAGCAGCGCACCTCGACGCTGATGATTC 2940
Qy 2941 AAGGAACGTCGCAACAGCCTCAACCGCTCTGTTCCGAAAGCCAAACGGAAGATTTCTC 3000
Db 2941 AAGGAACGTCGCAACAGCCTCAACCGCTCTGTTCCGAAAGCCAAACGGAAGATTTCTC 3000
Qy 3001 GAGCACCGTCCGCTTCGGCAACACCTCTGCGCTGATGATCGTGAATTTCTTACGGA 3060
Db 3001 GAGCACCGTCCGCTTCGGCAACACCTCTGCGCTGATGATCGTGAATTTCTTACGGA 3060
Qy 3061 CTGTCGAGGGCGCGAGACTTTGATCCGCTGCGCTGCGCAGATGTGCGCACCCCACTGCTTT 3120
Db 3061 CTGTCGAGGGCGCGAGACTTTGATCCGCTGCGCTGCGCAGATGTGCGCACCCCACTGCTTT 3120
Qy 3121 CGCTGATCGGATCTCTGAGCCAGAGATGAGGATGCGCAATGTTGTGGCCAAAGTTC 3180
Db 3121 CGCTGATCGGATCTCTGAGCCAGAGATGAGGATGCGCAATGTTGTGGCCAAAGTTC 3180
Qy 3181 AAGGGCAGATCCGCCCAATGCGCTGCGTGAACCGCTTGAAGTCTGTCAACCGCAACC 3240
Db 3181 AAGGGCAGATCCGCCCAATGCGCTGCGTGAACCGCTTGAAGTCTGTCAACCGCAACC 3240
Qy 3241 GCAGAAAGCAGATTCCTCCAAAGGGCCATTTGCTGCAACCATTCGCTGGTGGTTC 3300
Db 3241 GCAGAAAGCAGATTCCTCCAAAGGGCCATTTGCTGCAACCATTCGCTGGTGGTTC 3300
Qy 3301 ACTGTGATCTTCTGAGGATGATGAGTCAAGCTGAGATGAGTCAAGCTGAGATGAGTCAAG 3360
Db 3301 ACTGTGATCTTCTGAGGATGATGAGTCAAGCTGAGATGAGTCAAGCTGAGATGAGTCAAG 3360
Qy 3361 GCTATGAAGTGAAGCAACAATCACTGCTTCTGTCACGCAAGATTTGAACGGTGGT 3420
Db 3361 GCTATGAAGTGAAGCAACAATCACTGCTTCTGTCACGCAAGATTTGAACGGTGGT 3420
Qy 3421 GTTCTGCTGCAACGAAAGGTGGAAGGTGGGACCTTGAATCGTGGTTCCTAA 3474
Db 3421 GTTCTGCTGCAACGAAAGGTGGAAGGTGGGACCTTGAATCGTGGTTCCTAA 3474

RESULT 2

AX453603

LOCUS

DEFINITION

AX453603

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Corynebacterium glutamicum.

Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

Actinomycetales; Corynebacterineae; Corynebacteriaceae;

Corynebacterium.

REFERENCE

1

Hanke, P.D.

Feedback-resistant pyruvate carboxylase gene from corynebacterium

Patent: WO 0231158-A 3 18-APR-2002;

ARCHER-DANIELS-MIDLAND COMPANY (US)

FEATURES

source

1..3474

AX453603

Sequence 3 from Patent WO231158.

AX453603

AX453603.1 GI:21712843

Corynebacterium glutamicum.

Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

Actinomycetales; Corynebacterineae; Corynebacteriaceae;

Corynebacterium.

REFERENCE

1

Hanke, P.D.

Feedback-resistant pyruvate carboxylase gene from corynebacterium

Patent: WO 0231158-A 3 18-APR-2002;

ARCHER-DANIELS-MIDLAND COMPANY (US)

FEATURES

source

1..3474

Corynebacterium glutamicum.

Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

Actinomycetales; Corynebacterineae; Corynebacteriaceae;

Corynebacterium.

REFERENCE

1

Hanke, P.D.

Feedback-resistant pyruvate carboxylase gene from corynebacterium

Patent: WO 0231158-A 3 18-APR-2002;

ARCHER-DANIELS-MIDLAND COMPANY (US)

FEATURES

source

1..3474

/organism="Corynebacterium glutamicum"									
/db_xref="taxon:1718" 762 t									
BASE COUNT	749	a	1013	c	950	g	762	t	
ORIGIN									
Query Match	100.0%; Score 3474; DB 6; Length 3474;								
Best Local Similarity	100.0%; Pred. No. 0;								
Matches 3474; Conservative	0; Mismatches 0; Indels 0; Gaps 0;								
Qy	1	GTGACTGCTATCACCCCTTGGCGGTCTCTGTGTAAGGAATAATTAATCTAGTGTGCACT	60						
Db	1	GTGACTGCTATCACCCCTTGGCGGTCTCTGTGTAAGGAATAATTAATCTAGTGTGCACT	60						
Qy	61	CACACATCTTCAACGGTCTCCAGCATTTCAAAAAGATCTTGTAGCAAAACCGCGCGAATC	120						
Db	61	CACACATCTTCAACGGTCTCCAGCATTTCAAAAAGATCTTGTAGCAAAACCGCGCGAATC	120						
Qy	121	GCGGTCCGGTCTTCCGTGCAGCACTCGAAGCGTCTTTCGCTTCTGAAGCTGCGCAATGGTACT	180						
Db	121	GCGGTCCGGTCTTCCGTGCAGCACTCGAAGCGTCTTTCGCTTCTGAAGCTGCGCAATGGTACT	180						
Qy	181	CGTGAAGATCGGGATCATTCACCGCTCTTTCGCTTCTGAAGCTGCGCAATGGTACT	240						
Db	181	CGTGAAGATCGGGATCATTCACCGCTCTTTCGCTTCTGAAGCTGCGCAATGGTACT	240						
Qy	241	GAAGGCTCAAGTCAAGGCTGACATCGATGAATTAATCGGTGAGCTAAAAA	300						
Db	241	GAAGGCTCAAGTCAAGGCTGACATCGATGAATTAATCGGTGAGCTAAAAA	300						
Qy	301	GTTAAAGCAGATGCTATTATACCGGGATATGCTTCTGCTGAAAAATGCCAGCTGCC	360						
Db	301	GTTAAAGCAGATGCTATTATACCGGGATATGCTTCTGCTGAAAAATGCCAGCTGCC	360						
Qy	361	CGCGAGTCCGGGAAAAACGGCATTAATTTATGGCCCAACCCAGAGGTTCCTGATCTC	420						
Db	361	CGCGAGTCCGGGAAAAACGGCATTAATTTATGGCCCAACCCAGAGGTTCCTGATCTC	420						
Qy	421	ACCGGTGATAAGTCTGTCGGGTAAACCCCGGAAGAGGCTGCTGCCAGTTTGGCG	480						
Db	421	ACCGGTGATAAGTCTGTCGGGTAAACCCCGGAAGAGGCTGCTGCCAGTTTGGCG	480						
Qy	481	GAATCCACCCGAGCAAAACATGATGACATCGTTAAAGCGCTGAAGCCAGACTTAC	540						
Db	481	GAATCCACCCGAGCAAAACATGATGACATCGTTAAAGCGCTGAAGCCAGACTTAC	540						
Qy	541	CCCATCTTTGTAAGGCAGTTGCGGTGCGGAGCGGTATGCGCTTGTTCCTCA	600						
Db	541	CCCATCTTTGTAAGGCAGTTGCGGTGCGGAGCGGTATGCGCTTGTTCCTCA	600						
Qy	601	CTGTAGAGCTCCGCAAAATGGCAACAGAGCATCTGTGAAGCTGAAGCGGATTCGGC	660						
Db	601	CTGTAGAGCTCCGCAAAATGGCAACAGAGCATCTGTGAAGCTGAAGCGGATTCGGC	660						
Qy	661	GACGGTTGCGGTATATGTGCAAGCTGCTGTGATTAAACCCCGACACATTTGAAGTGCAGATC	720						
Db	661	GACGGTTGCGGTATATGTGCAAGCTGCTGTGATTAAACCCCGACACATTTGAAGTGCAGATC	720						
Qy	721	CTTGGCGATCGCACTGGAAGTTGTACACCTTTTATGAAGCTGACTGCTCACTGCAGCGT	780						
Db	721	CTTGGCGATCGCACTGGAAGTTGTACACCTTTTATGAAGCTGACTGCTCACTGCAGCGT	780						
Qy	781	CGTCAACAAAAGTTGTGAAAATTCGCGCAGACAGCATTTGGATCCAGAACTGCGTGAT	840						
Db	781	CGTCAACAAAAGTTGTGAAAATTCGCGCAGACAGCATTTGGATCCAGAACTGCGTGAT	840						
Qy	841	CGCATTTGTGCGGATCGACTGAAGTTCTGCCGCTCCATTTGGTTACCAAGGCGGGGAACC	900						
Db	841	CGCATTTGTGCGGATCGACTGAAGTTCTGCCGCTCCATTTGGTTACCAAGGCGGGGAACC	900						
Qy	901	GTGGAATCTTGTGATGAAAGGGAACCAAGCTTTTCATCGAAATGAACCCCAAGTATC	960						
Db	901	GTGGAATCTTGTGATGAAAGGGAACCAAGCTTTTCATCGAAATGAACCCCAAGTATC	960						
Qy	961	CAGGTTGAGCACACCGTGACTGAAGTACCGAGGTGGACCTGGTGAAGCGCGAGATG	1020						
Db	961	CAGGTTGAGCACACCGTGACTGAAGTACCGAGGTGGACCTGGTGAAGCGCGAGATG	1020						
Qy	1021	CGTTGGTCTGTGTGCAACCTTTGAAGGAATTTGGTCTGACCCCAAGATAAGATCAAGACC	1080						
Db	1021	CGTTGGTCTGTGTGCAACCTTTGAAGGAATTTGGTCTGACCCCAAGATAAGATCAAGACC	1080						
Qy	1081	CACGGTGCAGCACTGCAAGTCCGCTACCCAGCGGAGATCCAAACAACGGCTTCGGCCA	1140						
Db	1081	CACGGTGCAGCACTGCAAGTCCGCTACCCAGCGGAGATCCAAACAACGGCTTCGGCCA	1140						
Qy	1141	GATACCGGAAATATACCCGCTACCGCTACCCAGCGGAGCTGGCGTTTCGTCTTGAAGGT	1200						
Db	1141	GATACCGGAAATATACCCGCTACCGCTACCCAGCGGAGCTGGCGTTTCGTCTTGAAGGT	1200						
Qy	1201	GCAGTCAAGCTCGGTGGCGAAATCACCGCACACTTTTGACTCTCATCTGGTGAANAATGACC	1260						
Db	1201	GCAGTCAAGCTCGGTGGCGAAATCACCGCACACTTTTGACTCTCATCTGGTGAANAATGACC	1260						
Qy	1261	TGCGGTGGTCCGACTTTTGAACCTTGCTTCTGTCGACAGCGCGCTTGCTGAGTTTC	1320						
Db	1261	TGCGGTGGTCCGACTTTTGAACCTTGCTTCTGTCGTCGACAGCGCGCTTGCTGAGTTTC	1320						
Qy	1321	ACCGTGTCTGTGTGCAACCAACATTTGTTTCTGCTGCGGTGCGGGAAGAGGAC	1380						
Db	1321	ACCGTGTCTGTGTGCAACCAACATTTGTTTCTGCTGCGGTGCGGGAAGAGGAC	1380						
Qy	1381	TTTACTTCCAAAGCGCATCGCCACCGGATTTATCGGGATCACCCACACTCTTTCAGGCT	1440						
Db	1381	TTTACTTCCAAAGCGCATCGCCACCGGATTTATCGGGATCACCCACACTCTTTCAGGCT	1440						
Qy	1441	CCACCTCGGATGATGAGCAGGAGCGCATCTTGGATTACTTTGGCAGATGTCAACCGTGAAC	1500						
Db	1441	CCACCTCGGATGATGAGCAGGAGCGCATCTTGGATTACTTTGGCAGATGTCAACCGTGAAC	1500						
Qy	1501	AAGCCTCATGTGTGCGCTCCAAAGGATTTGACGACCAATCGATAAGCTGCCCAACATC	1560						
Db	1501	AAGCCTCATGTGTGCGCTCCAAAGGATTTGACGACCAATCGATAAGCTGCCCAACATC	1560						
Qy	1561	AAGGATCTGCCACTGCCACCGCTTCCCGTGACCGCTTGAAGCAGCTTTGGCCAGCGCG	1620						
Db	1561	AAGGATCTGCCACTGCCACCGCTTCCCGTGACCGCTTGAAGCAGCTTTGGCCAGCGCG	1620						
Qy	1621	TTTGTCTGTGATCTCCGTGAGCAGGAGCGCATCTGGCAGTTACTGTATACCACTTCGCGAT	1680						
Db	1621	TTTGTCTGTGATCTCCGTGAGCAGGAGCGCATCTGGCAGTTACTGTATACCACTTCGCGAT	1680						
Qy	1681	GCACACAGCTCTTGTGTGACCGGAGTCCGCTCATTTGCGACTGAAGCTGCGGCGAG	1740						
Db	1681	GCACACAGCTCTTGTGTGACCGGAGTCCGCTCATTTGCGACTGAAGCTTGGGCGAG	1740						
Qy	1741	GCCTGCGAAAAGCTGACTCTGTGAGCTTTTGTTCGTGGAAGCTTGGGGCGCGGACCTAC	1800						
Db	1741	GCCTGCGAAAAGCTGACTCTGTGAGCTTTTGTTCGTGGAAGCTTGGGGCGCGGACCTAC	1800						
Qy	1801	GATGTGGCGATCGCTTTCTCTTTGAGGATCCGTGGGACAGGCTCGACAGCTGCGCGAG	1860						
Db	1801	GATGTGGCGATCGCTTTCTCTTTGAGGATCCGTGGGACAGGCTCGACAGCTGCGCGAG	1860						
Qy	1861	GCAGTCCGGAATGTAAACATTTGATGATCTGCTTTCGCGCGCGCAACACCGTGGGATACACC	1920						
Db	1861	GCAGTCCGGAATGTAAACATTTGATGATCTGCTTTCGCGCGCGCAACACCGTGGGATACACC	1920						
Qy	1921	CGGTACCAGACTCCGTCTGCGCGCTTTTAAAGGAAGCTGCCAGCTCCGGCGTGGAC	1980						
Db	1921	CGGTACCAGACTCCGTCTGCGCGCTTTTAAAGGAAGCTGCCAGCTCCGGCGTGGAC	1980						
Qy	1981	ATCTTTCGCGATCTTTCGAGCGCTTAAAGAGCTCTCCACAGATGGTCCAGCAATCGACCA	2040						
Db	1981	ATCTTTCGCGATCTTTCGAGCGCTTAAAGAGCTCTCCACAGATGGTCCAGCAATCGACCA	2040						
Qy	2041	GTCTTGGAGACCAACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTTGGTGAATCTCTCT	2100						

Db 2041 GTCTGGAGACCAACACCGCGGTAGCCAGGTGGCTATGGCTTATTTCTGGTGAATCTCTCT 2100
Qy
Db 2101 GATCCAAATGAAAGCTCTACACCCCTGGATTAACCTTAAGAGATGGCAGAGAGATCGTC 2160
Db 2101 GATCCAAATGAAAGCTCTACACCCCTGGATTAACCTTAAGAGATGGCAGAGAGATCGTC 2160
Qy 2161 AAGTCTGGCGCTCACATTTCTGGCCATTAAGAGATATGGCTGGTCTGCTTCGGCCAGCTGGC 2220
Db 2161 AAGTCTGGCGCTCACATTTCTGGCCATTAAGAGATATGGCTGGTCTGCTTCGGCCAGCTGGC 2220
Qy 2221 GTAACCAAGCTGGTCACCGCACTCGCGCTGAAATTCGATCTGCCAGTSCAGTGCACACC 2280
Db 2221 GTNACCAAGCTGGTCACCGCACTCGCGCTGAAATTCGATCTGCCAGTSCAGTGCACACC 2280
Qy 2281 CACGACACTCGGGGTGGCCAGTTGGCTACTACTTTGTCTGCAGCTCAAGCTGGTGCAGAT 2340
Db 2281 CACGACACTCGGGGTGGCCAGTTGGCTACTACTTTGTCTGCAGCTCAAGCTGGTGCAGAT 2340
Qy 2341 GCTGTTAGCGTGTCTCGGCACCACTGTCTGGCACCACTCTCCAGCCATCCCTGTCTGCC 2400
Db 2341 GCTGTTAGCGTGTCTCGGCACCACTGTCTGGCACCACTCTCCAGCCATCCCTGTCTGCC 2400
Qy 2401 ATTGTTGCTGCAATTCGCGCACACCCGTCGCGATACCGGTTTGGAGCTTCGAGCTCGAGCTGTTCT 2460
Db 2401 ATTGTTGCTGCAATTCGCGCACACCCGTCGCGATACCGGTTTGGAGCTTCGAGCTCGAGCTGTTCT 2460
Qy 2461 GACCTCGAGCCGTACTGGGAAGCTGTGCGCGAGCTGTACTCTGCCATTTAGCTTGGAAACC 2520
Db 2461 GACCTCGAGCCGTACTGGGAAGCTGTGCGCGAGCTGTACTCTGCCATTTAGCTTGGAAACC 2520
Qy 2521 CCAGGCCCAACCGTGGCTGTACCGGCACGAAATCCAGCGGAGACAGTTGTCACACCTG 2580
Db 2521 CCAGGCCCAACCGTGGCTGTACCGGCACGAAATCCAGCGGAGACAGTTGTCACACCTG 2580
Qy 2581 CGTGACAGGCCACCGCACTGGCTGTCTGATCGCTTCGAGCTCATCGAAGCAACTAC 2640
Db 2581 CGTGACAGGCCACCGCACTGGCTGTCTGATCGCTTCGAGCTCATCGAAGCAACTAC 2640
Qy 2641 GCAGCCGTTAATGAGATGCTGGGAGCGCCCAACCAAGGTCAACCCATCTCCCAAGGTTGTT 2700
Db 2641 GCAGCCGTTAATGAGATGCTGGGAGCGCCCAACCAAGGTCAACCCATCTCCCAAGGTTGTT 2700
Qy 2701 GGCAGCTCGCACTCCACCTGGTGGTGGCGGTGTAGATCCAGAGACTTTGCTGCAGAC 2760
Db 2701 GGCAGCTCGCACTCCACCTGGTGGTGGCGGTGTAGATCCAGAGACTTTGCTGCAGAC 2760
Qy 2761 CCACAAAGTACGACATCCAGACTCTGTATCGGTTCTCGCGGCGAGCTTGGTAAC 2820
Db 2761 CCACAAAGTACGACATCCAGACTCTGTATCGGTTCTCGCGGCGAGCTTGGTAAC 2820
Qy 2821 CCTCCAGTGGCTGGCCAGAACCACTGGCACCCGCGCACTGGAAAGGCGGCTCCGAAGGC 2880
Db 2821 CCTCCAGTGGCTGGCCAGAACCACTGGCACCCGCGCACTGGAAAGGCGGCTCCGAAGGC 2880
Qy 2881 AAGGCACTCTGACGGAAGTTCTGAGGAAGAGCAGGCGCACTTCGACGCTGATGATTCC 2940
Db 2881 AAGGCACTCTGACGGAAGTTCTGAGGAAGAGCAGGCGCACTTCGACGCTGATGATTCC 2940
Qy 2941 AAGGCACTCTGACGGAAGTTCTGAGGAAGAGCAGGCGCACTTCGACGCTGATGATTCC 3000
Db 2941 AAGGCACTCTGACGGAAGTTCTGAGGAAGAGCAGGCGCACTTCGACGCTGATGATTCC 3000
Qy 3001 GAGCACCGTCCCGCTTCGGCAACACCTCTCGCTGGATGATCGTGAATTTCTTACGGA 3060
Db 3001 GAGCACCGTCCCGCTTCGGCAACACCTCTCGCTGGATGATCGTGAATTTCTTACGGA 3060
Qy 3061 CTGTGAGGCGCCGAGACTTTGATCGCTGCCAGATGTGCGCACCCCACTGCTTTGTT 3120
Db 3061 CTGTGAGGCGCCGAGACTTTGATCGCTGCCAGATGTGCGCACCCCACTGCTTTGTT 3120
Qy 3121 CGCCTGATGCGATCTCTGACCCAGACGATTAAGGGTATGCGCAATGTTGTGGCCCAACGTC 3180

Db 3121 CGCCTGATGCGATCTCTGAGCCAGACGATTAAGGGTATGCGCAATGTTGTGGCCCAACGTC 3180
Qy 3181 AACGGCCAGATCCGCCCAATCGGTGTCGTGACCGCTCCGTTGAGTCTGTACCGCAACC 3240
Db 3181 AACGGCCAGATCCGCCCAATCGGTGTCGTGACCGCTCCGTTGAGTCTGTACCGCAACC 3240
Qy 3241 GCAGAAAAGGAGATTTCTTCCAAACAGGGCCATGTTCTGTCGACCAATTCGCTGGTGTGTC 3300
Db 3241 GCAGAAAAGGAGATTTCTTCCAAACAGGGCCATGTTCTGTCGACCAATTCGCTGGTGTGTC 3300
Qy 3301 ACTGTGATCTGTGCTGAAGGTGATGAGTCAAGGCTCGAGATGCAAGTGCATCATCGAG 3360
Db 3301 ACTGTGATCTGTGCTGAAGGTGATGAGTCAAGGCTCGAGATGCAAGTGCATCATCGAG 3360
Qy 3361 GCTATGAAGTGAAGCAACAATCACTGCTTCTGTTGACCGCAAGATTTGAACGCGTTGTG 3420
Db 3361 GCTATGAAGTGAAGCAACAATCACTGCTTCTGTTGACCGCAAGATTTGAACGCGTTGTG 3420
Qy 3421 GTTCTCTGCTCAACGAAAGGTGGAAGGTGGGAGCTTGTGATCGTCTGCTTTCCTAA 3474
Db 3421 GTTCTCTGCTCAACGAAAGGTGGAAGGTGGGAGCTTGTGATCGTCTGCTTTCCTAA 3474

RESULT 3
AR123974 LOCUS AR123974 3621 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6171833.
ACCESSION AR123974
VERSION AR123974.1 GI:14109335
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3621)
AUTHORS Sinek, A.J., Lessard, P.A. and Willis L.B.
TITLE Pyruvate carboxylase from corynebacterium glutamicum
JOURNAL Patent: US 6171833-A 1 09-JAN-2001;
FEATURES Location/Qualifiers
source 1..3621
BASE COUNT 782 a 1037 c 1002 g 800 t
ORIGIN

Query Match 97.8%; Score 3398.8; DB 6; Length 3621;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3427; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 1 GTGACTGCTATCACCCCTTGGCGGTCTCTTGTGAAAGGAATAATTACTTAGTGTGCACT 60
Db 148 GTGACTGCTATCACCCCTTGGCGGTCTCTTGTGAAAGGAATAATTACTTAGTGTGCACT 207
Qy 61 CACACATCTTCAACGCTTCCAGCATTCAAAAGATCTTGGTAGCAAAACCGCGGGAATC 120
Db 208 CACACATCTTCAACGCTTCCAGCATTCAAAAGATCTTGGTAGCAAAACCGCGGGAATC 267
Qy 121 GCGGTCCGTCTTTCGCGTGCAGCACTCGAAACCGGTGACGCGTAGCTATTATACCCC 180
Db 268 GCGGTCCGTCTTTCGCGTGCAGCACTCGAAACCGGTGACGCGTAGCTATTATACCCC 327
Qy 181 CGTGAGATCGGGATCATTTCCACCGCTCTTTGCTTCTGAAAGCTGTCGGAATGGTACT 240
Db 328 CGTGAGATCGGGATCATTTCCACCGCTCTTTGCTTCTGAAAGCTGTCGGAATGGTACT 387
Qy 241 GAAGGCTCACAGTCAAGGCTACCTGGACATCGATGAATTTATCGTGAGCTTAAAAA 300
Db 388 GAAGGCTCACAGTCAAGGCTACCTGGACATCGATGAATTTATCGTGAGCTTAAAAA 447
Qy 301 GTTAAAGCAGATGCTATTATACCGGGATATGGCTTCTGCTGAAATGCCAGCTTGC 360
Db 448 GTTAAAGCAGATGCTATTATACCGGGATATGGCTTCTGCTGAAATGCCAGCTTGC 507
Qy 361 CGGAGTGGCGGAAACCGCATTTATTTGTTGGCCCAACCCCGAGGTTCTTGATCTC 420

Db 508 CGCAGTGTGCGGAAAAAGGCAATTACTTTTATTTGGCCCAACCCACAGAGGTTCCTTGATCTC 567
Qy 421 ACCGGTGAATAAGTCTCGTGGGTAAACCCGCGCGAAGAGCTGCTGTGCGAGTTTGGCG 480
Db 568 ACCGGTGAATAAGTCTCGCGCGGTAAACCCGCGCGAAGAGCTGCTGTGCGAGTTTGGCG 627
Qy 481 GAATCCACCCCGAGCAAAAAACATCGATGACATCGTTAAAGAGCGCTGAAGGCCAGACTTAC 540
Db 628 GAATCCACCCCGAGCAAAAAACATCGATGAGATCGTTAAAGAGCGCTGAAGGCCAGACTTAC 687
Qy 541 CCCATCTTTGTAAGGCGAGTTGCGGTGTCGGCGGACGCGGTATGCGCTTGTGTTCTTCA 600
Db 688 CCCATCTTTGTAAGGCGAGTTGCGGTGTCGGCGGACGCGGTATGCGCTTGTGTTCTTCA 747
Qy 601 CTTGATGAGCTCGCGMAATTTGGCAACAGAGCATCTGTAAGCTGAAGCGCATTTGGC 660
Db 748 CTTGATGAGCTTCGCAAAATTTAGCAACAGAGCATCTGTAAGCTGAAGCGCATTTGGC 807
Qy 661 GACGGTTCGGTATATGTGGAACGTGCTGTGATTAACCCCGACGACATTTGAAGTGCAGATC 720
Db 808 GATGGCGCGGTATATGTGGAACGTGCTGTGATTAACCCCTCAGCATATTGAAGTGCAGATC 867
Qy 721 CTTGGCGATCGCACTGGAGAGTTGTACACCTTTATGAACGTGACTGCTCCTCAGCGGT 780
Db 868 CTTGGCGATCACACTGGAGAGTTGTACACCTTTATGAACGTGACTGCTCCTCAGCGGT 927
Qy 781 CGTCACAAAAAGTTGTCGAAATTCGCGCAGCACAGATTTGGATCCAGAACTGCGTGAT 840
Db 928 CGTCACAAAAAGTTGTCGAAATTCGCGCAGCACAGATTTGGATCCAGAACTGCGTGAT 987
Qy 841 CGATTTGTGGATGTCAGTAAGTTCTGCGCTCCATTTGTTTACAGGCGCGGGAAC 900
Db 988 CGATTTGTGGATGTCAGTAAGTTCTGCGCTCCATTTGTTTACAGGCGCGGGAAC 1047
Qy 901 GTGGAAATCTTGGTGATGAAGAGGCAACAGTTTTCATTCGAAATGACCCAGTATC 960
Db 1048 GTGGAAATCTTGGTGATGAAGAGGCAACAGTTTTCATTCGAAATGACCCAGTATC 1107
Qy 961 CAGGTTGAGCACACCGTGACTGAGAGTCAACAGGTGACCTGTGTGAAGGCGCAGATG 1020
Db 1108 CAGGTTGAGCACACCGTGACTGAGAGTCAACAGGTGACCTGTGTGAAGGCGCAGATG 1167
Qy 1021 CGTTGGCTGTGGTGCAACCTTTGAAGGAAATTTGGGTCTGACCCCAAGATAAGATCAAGACC 1080
Db 1168 CGTTGGCTGTGGTGCAACCTTTGAAGGAAATTTGGGTCTGACCCCAAGATAAGATCAAGACC 1227
Qy 1081 CACGGTGCAAGCTGCAAGTCCGCAATCAACGGAAGATCCAAACAAACGCGTTTCGCCCCA 1140
Db 1228 CACGGTGCAAGCTGCAAGTCCGCAATCAACGGAAGATCCAAACAAACGCGTTTCGCCCCA 1287
Qy 1141 GATACCGGAACATCACCGCGTACCGCTCACAGCGGAGCTGGCGTTCTGTGACGGT 1200
Db 1288 GATACCGGAACATCACCGCGTACCGCTCACAGCGGAGCTGGCGTTCTGTGACGGT 1347
Qy 1201 GCAGCTCAGCTCGGTGGGAAATCACCGCACACTTTGACTCCATGCTGGTGAAATGACC 1260
Db 1348 GCAGCTCAGCTCGGTGGGAAATCACCGCACACTTTGACTCCATGCTGGTGAAATGACC 1407
Qy 1261 TGGCGGTGTCGACCTTGAAGTCTGCTGCTGTCAGAGCGCGTGGCTGAGTTC 1320
Db 1408 TGGCGGTGTCGACCTTGAAGTCTGCTGCTGTCAGAGCGCGTGGCTGAGTTC 1467
Qy 1321 ACCGTGCTGGTGTGCAACCAATTTGGTTCTTGGCTGCTGCTGCGGGAAGAGAC 1380
Db 1468 ACCGTGCTGGTGTGCAACCAATTTGGTTCTTGGCTGCTGCTGCGGGAAGAGAC 1527
Qy 1381 TTCACTTCCAAAGCGCATCGCCACCGGATTTATCGCGCATACCCACACTCTTCAAGGT 1440
Db 1528 TTCACTTCCAAAGCGCATCGCCACCGGATTTATCGCGCATACCCCGCACTCTTCAAGGT 1587
Qy 1441 CCACCTGCGGATGATGAGAGGAGCGATCTCTGGATTTACTTGGCAGATGTCAACCGTGAAC 1500
Db 1588 CCACCTGCGGATGATGAGAGGAGCGATCTCTGGATTTACTTGGCAGATGTCAACCGTGAAC 1647

Qy 1501 AAGCCTCATGTGTGCGTCCAAAGGATGTGCGAGCACAAATCGATAAGCTGCCCAACATC 1560
Db 1648 AAGCCTCATGTGTGCGTCCAAAGGATGTGCGAGCTCTATCGATAAGCTGCCCAACATC 1707
Qy 1561 AAGGATCTGCCACTGCCACGCGGTTCGCGTGACCGCTGAAGCAGCTTTGGGCCAGCGCG 1620
Db 1708 AAGGATCTGCCACTGCCACGCGGTTCGCGTGACCGCTGAAGCAGCTTTGGGCCAGCGCG 1767
Qy 1621 TTTGCTCGTGTATCTCCGTGAGCAGGACGCACTGGCAGTTACTGTATACCACTTCCGCGAT 1680
Db 1768 TTTGCTCGTGTATCTCCGTGAGCAGGACGCACTGGCAGTTACTGTATACCACTTCCGCGAT 1827
Qy 1681 GCACACCACTCTTTGCTTGCACCCGAGTCCGCTCATTCGCACTGAAGCTGCGSCAGAG 1740
Db 1828 GCACACCACTCTTTGCTTGCACCCGAGTCCGCTCATTCGCACTGAAGCTTGGSCAGAG 1887
Qy 1741 GCGGTGCAAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGCCTGGGCGGCGGACCTAC 1800
Db 1888 GCGGTGCAAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGCCTGGGCGGCGGACCTAC 1947
Qy 1801 GATGTGGCGATGCGTTCCTTTGAGGATCCGTGGGACAGCTCGACGAGCTGGCGAG 1860
Db 1948 GATGTGGCGATGCGTTCCTTTGAGGATCCGTGGGACAGCTCGACGAGCTGGCGAG 2007
Qy 1861 GCGATGCCGAATGTAAACATTCAGATGCTCTTCGCGGCCGCAACACCGTGGGATACACC 1920
Db 2008 GCGATGCCGAATGTAAACATTCAGATGCTCTTCGCGGCCGCAACACCGTGGGATACACC 2067
Qy 1921 CCGTACCCAGACTCCGTCTGCGCGCGTTCGTTAAGGAAAGCTGCCAGCTCCGGCGTGGAC 1980
Db 2068 CCGTACCCAGACTCCGTCTGCGCGCGTTCGTTAAGGAAAGCTGCCAGCTCCGGCGTGGAC 2127
Qy 1981 ATCTTCCGCATCTTCGACGCGCTTAAAGAGCTCTCCAGATCGTCCAGCAATCCAGCA 2040
Db 2128 ATCTTCCGCATCTTCGACGCGCTTAAAGAGCTCTCCAGATCGTCCAGCAATCCAGCA 2187
Qy 2041 CTCCTGAGACCAACACCGGTAGCGAGGTGGCTATCGCTTATTCGTGATCTCTCT 2100
Db 2188 GTCTTGGAGACCAACACCGGTAGCGAGGTGGCTATCGCTTATTCGTGATCTCTCT 2247
Qy 2101 GATCAAAATGAAAAGCTCTACACCTGGATTACTCTAAAGATGGCAGAGATCTC 2160
Db 2248 GATCAAAATGAAAAGCTCTACACCTGGATTACTCTAAAGATGGCAGAGATCTC 2307
Qy 2161 AAGTCTGGCGTCTCAATCTTGGGCAATTAAGGATATGGCTGGTCTGCTGCGCCAGCTGG 2220
Db 2308 AAGTCTGGCGTCTCAATCTTGGGCAATTAAGGATATGGCTGGTCTGCTGCGCCAGCTGG 2367
Qy 2221 GTAACCAAGCTGGTCAACGCACTGCGCGCTGAATTCGATCTGCCAGTGCACTGCAACC 2280
Db 2368 GTAACCAAGCTGGTCAACGCACTGCGCGCTGAATTCGATCTGCCAGTGCACTGCAACC 2427
Qy 2281 CACGACACTGCGGGTGGCCAGTTCGCTACCTTTGCTGCGAGCTCAAGCTGGTGCAGAT 2340
Db 2428 CACGACACTGCGGGTGGCCAGTTCGCTACCTTTGCTGCGAGCTCAAGCTGGTGCAGAT 2487
Qy 2341 GCTGTTGACGGTGTCTCCGACCACTGTGCGCACACCTCCAGCACTCCCTGTCTGCC 2400
Db 2488 GCTGTTGACGGTGTCTCCGACCACTGTGCGCACACCTCCAGCACTCCCTGTCTGCC 2547
Qy 2401 ATTGTTGCTGCAATTCGCGCACACCGCTGCGATACCGGTTTGAGCTCGAGGCTCTTCT 2460
Db 2548 ATTGTTGCTGCAATTCGCGCACACCGCTGCGATACCGGTTTGAGCTCGAGGCTCTTCT 2607
Qy 2461 GACCTCGAGCGCTACTGCGGAAGCTGTGCGGGAATGTACTGCGCAATTTGAGTCTGGAACC 2520
Db 2608 GACCTCGAGCGCTACTGCGGAAGAGTGTGCGGGAATGTACTGCGCAATTTGAGTCTGGAACC 2667
Qy 2521 CCAGGCCCAACCGGTGCGGTCTACCGCCACGAAATCCAGGCGGAGCAGTTGTCCAACTGT 2580
Db 2668 CCAGGCCCAACCGGTGCGGTCTACCGCCACGAAATCCAGGCGGAGCAGTTGTGTCCAACTGT 2727

Qy	2581	CGTGACAGGCGACCGCACTGGGCTTGCTGATCGCTCGAGCTCATCGAAGCAACTAC	2640
Db	2728	CGTGACAGGCGACCGCACTGGGCTTGCGGATCGTTTCGAACTCATCGAAGCAACTAC	2787
Qy	2641	GCAGCCGTTAATGAGATGCTCGGACGCCCAACCAAGGTCAACCCATCTCCCAAGTGTGT	2700
Db	2788	GCAGCCGTTAATGAGATGCTCGGACGCCCAACCAAGGTCAACCCATCTCCCAAGTGTGT	2847
Qy	2701	GGGACCTCGCATCTCCACCTGGTGGTGCGGGGTAGATCCAGCAGACTTTGCTGCAGAC	2760
Db	2848	GGGACCTCGCATCTCCACCTGGTGGTGCGGGGTAGATCCAGCAGACTTTGCTGCAGAT	2907
Qy	2761	CCACAAAGTAGACATCCCAAGCTCTGTATCGGTTCCTCGCGGCGAGCTTGTAAAC	2820
Db	2908	CCACAAAGTAGACATCCCAAGCTCTGTATCGGTTCCTCGCGGCGAGCTTGTGTAAAC	2967
Qy	2821	CCTCAGGTGCTGGCCAGAACCACTGCGCACCCCGCGCACTCGAAGGCGCTCCGAAGGC	2880
Db	2968	CCTCAGGTGCTGGCCAGAGCCACTGCGCACCCCGCGCACTGGAAGGCGCTCCGAAGGC	3027
Qy	2881	AAGGACCTCTGACGGAGATTCCTGAGGAAGAGAGAGGCGCACTCGAGCTGATGATTC	2940
Db	3028	AAGGACCTCTGACGGAGATTCCTGAGGAAGAGAGAGGCGCACTCGAGCTGATGATTC	3087
Qy	2941	AAGGAAGCTGCAACAGCCTCAACGGCTGCTGTCTCCGAAGCCCAACCGAAGATTCCTC	3000
Db	3088	AAGGAAGCTGCAATAGCTCAACGGCTGCTGTCTCCGAAGCCCAACCGAAGATTCCTC	3147
Qy	3001	GAGCACCGTCCCGCTTCGGCAACACCTCTCGCTGATGATCGTGAATCTTCTACGGA	3060
Db	3148	GAGCACCGTCCCGCTTCGGCAACACCTCTCGCTGATGATCGTGAATCTTCTACGCG	3207
Qy	3061	CTGTGAGGCGCCGAGACTTTGATCGCTGCCAGATGTCGCAACCCCACTGCTTGT	3120
Db	3208	CTGTGAGGCGCCGAGACTTTGATCGCTGCCAGATGTCGCAACCCCACTGCTTGT	3267
Qy	3121	CGCTGATGCGATCTCTGAGCAGAGATGAAGGGTATGCCAATGTTGGCCAAAGTC	3180
Db	3268	CGCTGATGCGATCTCTGAGCAGAGATGAAGGGTATGCCAATGTTGGCCAAAGTC	3327
Qy	3181	AACGGCAGATCCGCCCAATCGGTGCGTGCACCGCTCGTGTGAGTCTGTCAACCGCAAC	3240
Db	3328	AACGGCAGATCCGCCCAATCGGTGCGTGCACCGCTCGTGTGAGTCTGTCAACCGCAAC	3387
Qy	3241	GCAGAAAGGCGAGATTCCTCCAAAGGCGCAATGTTGTCGACCACTTCGCTGTTGTC	3300
Db	3388	GCAGAAAGGCGAGATTCCTCCAAAGGCGCAATGTTGTCGACCACTTCGCTGTTGTC	3447
Qy	3301	ACTGTGACTGTTGCTGAAGTGTAGGTCAAGGCTGAGATGAGTCCGCAATCATCGAG	3360
Db	3448	ACCGTGACTGTTGCTGAAGTGTAGGTCAAGGCTGAGATGAGTCCGCAATCATCGAG	3507
Qy	3361	GCTATGAAGTGAAGCAACAATCACTGCTTCTGTGACGGCAAGATTGAACGGTGTG	3420
Db	3508	GCTATGAAGTGAAGCAACAATCACTGCTTCTGTGACGGCAAAATCGATCGGCTGTG	3567
Qy	3421	GTTCTCTGCTCAACGAAGGTGAAGGTGGGACTTGTGATCGCTGCTTTCCTAA	3474
Db	3568	GTTCTCTGCTCAACGAAGGTGAAGGTGGGACTTGTGATCGCTGCTTTCCTAA	3621
RESULT 4	AF038548	3637 bp	DNA linear BCT 25-AUG-2000
LOCUS	Corynebacterium glutamicum pyruvate carboxylase (pyc) gene, complete cds.		
DEFINITION	AF038548		
ACCESSION	AF038548		
VERSION	AF038548.1		
KEYWORDS	GI:2708716		
SOURCE	Corynebacterium glutamicum.		
ORGANISM	Corynebacterium glutamicum		
	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;		
	Actinomycetales; Corynebacterineae; Corynebacteriaceae;		

REFERENCE	Corynebacterium.
AUTHORS	1 (bases 1 to 3637) Koffas,M.A., Ramamoorthi,R., Pine,W.A., Sinskey,A.J. and Stephanopoulos,G.
TITLE	Sequence of the Corynebacterium glutamicum pyruvate carboxylase gene
JOURNAL	Appl. Microbiol. Biotechnol. 50 (3), 346-352 (1998)
MEDLINE	99019028
PUBMED	9802220
REFERENCE	2 (bases 1 to 3637) Koffas,M.A.G., Ramamoorthi,R., Pine,W.A., Sinskey,A.J. and Stephanopoulos,G.
AUTHORS	Direct Submission
TITLE	Submitted (14-DEC-1997) Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
JOURNAL	
FEATURES	Location/Qualifiers
source	1..3637 /organism="Corynebacterium glutamicum" /strain="21253" /db_xref="taxon:1718" 199..3621 /gene="pyc" 199..3621 /gene="pyc" /codon_start=1 /transl_table=11 /product="pyruvate carboxylase" /protein_id="AAB92588.1" /db_xref="GI:2708717" /translation="MSTHTSSLTLPAPFKTLVANRGEIARAPRAALETGAATVAIYPR EDRGSFHSFASEAVRIGTESPVAAYLDIDEIGAAKKVADAIYPGVFLSENAQL ARECANGITFTGPTPEVLDTLGSRAVTAAGKLPVLAESTPSKNDIEIVKSAEG QTEYPIFKAVAGGGRMFVAPDRELKLAETAREAAAGDGVAVYVERAVINPQH ISVOILGDHTGEVHLVYRDCSILORHOKVETIAPAHLDPDLRICALADAVKFCRSI GYOGAGTVEFLVDEKGNHVFIEWNPRILOVEHTVTEVSDLVKAOMLAAGATLKEI GLTQDKIKTHGAALQCRITTEDPNNGFRPDGTIITAYRSPGAGVLDGAAQCGEIT AFDMSLVKMTCRGDFETAVARAQALAEFTVSGVATNIGFLRLLREEDFTSKRIA TGFIDAPHLQLQAPPADDEQRIIDYLDADVTYKPHGVKPDVAAPIDKLPNKLDLP PRGSRDLQKLGPAAFARDLREQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVA KLTPILLSVEAMGGATYDVMRFLPEDWDLDELREAMPNVNIOMLLRGNTVGTVP YDPSVCRAVKEAASSGVDFRIFDALNDVSQMRPAIDAVLENTAVAVMAVYSGDL SPNEKLYTLIDYIKWAEIIVKGAHILAIKDMAGLRPAATKLVLTURREFDLPHV VHTHTAGGLATYFAAAGADAVDGAAPISGTSQPSLSAIVAAFAHTRRDITGLS LEAVSDLEPYWEAVRGLYLPFFSGTGPGRVYRHEIPGGQLSNRAQATLGLDRF ELIEDNYAAVNEMLGRPTKVTSPKVVGDALHLVAGVDPAADPOKDYIDPSVI AFRLGELNPPCGWPEPLTRALEGRSECKAPLVEPBEEOAHLDDADDKERNNSLNR LLPFPKTEEFLEHRRRFGNTSALDDBREFPYGLVEGRETLIRLPDVTPLLRDLDAISE PDDKGRNVVAVNQIRPMRVDRDSVESVTAEKADSSNKGHVAAPFAGVVTVA EGDEVKAGDAVAIIIEAMKMEATITASVDGKIDRVVVVPAATKVEGGDLIVVVS"
BASE COUNT	788 a 1040 c 1004 g 805 t
ORIGIN	
Query Match	97.8%; Score 3398.8; DB 1; Length 3637;
Best Local Similarity	98.6%; Pred. No. 0;
Matches 3427; Conservative	0; Mismatches 47; Indels 0; Gaps 0;
Qy	1 GTGACTGTATCACCCCTTGGCGGTCTCTTCTGAAAGGAATAATTAATCTAGTGTGCACT 60
Db	
Db	148 GTGACTGTATCACCCCTTGGCGGTCTCTTCTGAAAGGAATAATTAATCTAGTGTGCACT 207
Qy	61 CACACATCTTCAACCGCTTCCAGCATTCACAAAGATCTTGGTAGCAACCGCGCGGAATC 120
Db	
Db	208 CACACATCTTCAACCGCTTCCAGCATTCACAAAGATCTTGGTAGCAACCGCGCGGAATC 267
Qy	121 GCGGTGCGTCTTCCGTGAGCACTCGAAACCGGTGCGAGCCAGGTAGTATTATACCC 180
Db	
Db	268 GCGGTGCGTCTTCCGTGAGCACTCGAAACCGGTGCGAGCCAGGTAGTATTATACCC 327
Qy	181 CGTGAAGATCGGGATCATTCACCGCTCTTTTGTCTTCTGAAGCTGTCCGCAATGGTACT 240
Db	
Db	328 CGTGAAGATCGGGATCATTCACCGCTCTTTTGTCTTCTGAAGCTGTCCGCAATGGTACT 387

QY 241 GAAGGCTCACCAGTCAAGCGTACCTGGACATCGATGAATTTATCGGTGACGCTAAAAAA 300
Db 388 GAAGGCTCACCAGTCAAGCGTACCTGGACATCGATGAATTTATCGGTGACGCTAAAAAA 447
QY 301 GTTAAAGCAGATGCTATTATACCCGGGATATACGCTTCTGCTGAAAATGCCAGCTTGCC 360
Db 448 GTTAAAGCAGATGCCATTATACCCGGGATACGGCTTCTGCTGAAAATGCCAGCTTGCC 507
QY 361 CGGAGTGGCGGAAACGGCAATTACTTTTATGGCCCAACCCAGAGGTTCTTGATCTC 420
Db 508 CGGAGTGGCGGAAACGGCAATTACTTTTATGGCCCAACCCAGAGGTTCTTGATCTC 567
QY 421 ACCGGTGATAAGTCTCGTGGGTAAACCCCGGAAAGGCTGCTGCGCAGTTTGGCG 480
Db 568 ACCGGTGATAAGTCTCGTGGGTAAACCCCGGAAAGGCTGCTGCGCAGTTTGGCG 627
QY 481 GAATCACCACCGAGCAAAACATCGATCACATCGTTAAAGCGCTGAAGCCAGACTTAC 540
Db 628 GAATCACCACCGAGCAAAACATCGATCAGATCGTTAAAGCGCTGAAGCCAGACTTAC 687
QY 541 CCGATCTTTGTAAAGCAGTTCCGGTGGTGGGAGCGCGTATGCGCTTTGTTCTCA 600
Db 688 CCGATCTTTGTAAAGCAGTTCCGGTGGTGGGAGCGCGTATGCGCTTTGTTCTCA 747
QY 601 CTTGATGAGCTCGCMAATTTGGCAACAGAGCATCTCGTAAGCTGAAGCGCATTCGCG 660
Db 748 CTTGATGAGCTCGCMAATTTAGCAACAGAGCATCTCGTAAGCTGAAGCGCATTCGCG 807
QY 661 GACGGTTCCGTATATGCGAAGCTGCTGTGATTAACCCCGCAGCACATGAAGTGCAGATC 720
Db 808 GATGGCGCGGTATATGCGAAGCTGCTGTGATTAACCCCTCAGCATATGAAGTGCAGATC 867
QY 721 CTTGGCGGATCGCATGAGAGGTTGTACACCTTTATGAAGCTGACTGCTCACTGCAAGCT 780
Db 868 CTTGGCGGATCGCATGAGAGGTTGTACACCTTTATGAAGCTGACTGCTCACTGCAAGCT 927
QY 781 CGTCACCAAAAGTTGTGAAAATTCGCGCAGCACAGCATTTGGATCAGAACTTGGGTGAT 840
Db 928 CGTCACCAAAAGTTGTGAAAATTCGCGCAGCACAGCATTTGGATCAGAACTTGGGTGAT 987
QY 841 CGATTGTGCGGATGACGTAAGTTCTGCGCTCCATTTGTTACCAAGCGCGGGAACC 900
Db 988 CGCATTTGTGCGGATGACGTAAGTTCTGCGCTCCATTTGTTACCAAGCGCGGGAACC 1047
QY 901 GTGGAATTTCTGGTGCATGAAGGCAACCAAGTTTTCATCGAATGACCCAGCTATC 960
Db 1048 GTGGAATTTCTGGTGCATGAAGGCAACCAAGTTTTCATCGAATGACCCAGCTATC 1107
QY 961 CAGTTGAGCACACCGTGAAGTCAAGAGTCAAGAGTGGACCTGGTGAAGCGCGAGATG 1020
Db 1108 CAGTTGAGCACACCGTGAAGTCAAGAGTCAAGAGTGGACCTGGTGAAGCGCGAGATG 1167
QY 1021 CGTTGGCTGCTGGTGCAACCTTTGAAGGAATTTGGGTCTGACCCAAAGATAAGATCAAGACC 1080
Db 1168 CGTTGGCTGCTGGTGCAACCTTTGAAGGAATTTGGGTCTGACCCAAAGATAAGATCAAGACC 1227
QY 1081 CACGGTGACGACTGCAAGTCCGCGATCAACCGGAAGATCCAAACACAGGCTTTCGCCCCA 1140
Db 1228 CACGGTGACGACTGCAAGTCCGCGATCAACCGGAAGATCCAAACACAGGCTTTCGCCCCA 1287
QY 1141 GATACCGGAACACTACACCGGTACCGCTCACAGCGGAGCTGGCGTTCTGTTGACGGT 1200
Db 1288 GATACCGGAACACTACACCGGTACCGCTCACAGCGGAGCTGGCGTTCTGTTGACGGT 1347
QY 1201 GCAGCTCAGCTCGGTGGGAAATCAACCGCACACTTTGACTCCATGCTGGTGAATAAGACC 1260
Db 1348 GCAGCTCAGCTCGGTGGGAAATCAACCGCACACTTTGACTCCATGCTGGTGAATAAGACC 1407
QY 1261 TGGCGGTTCGACATTGAAATCTGTTGCTGTGCACAGCGCGGTTGGCTGAGTTTC 1320
Db 1408 TGGCGGTTCGACATTGAAATCTGTTGCTGTGCACAGCGCGGTTGGCTGAGTTTC 1467
QY 1321 ACCGTGCTGGTGTGCAACCAACATTTGGTTCTTGGCGTGGTGTGCGGGAAGAGAC 1380

Db 1468 ACCGTGCTGGTGTGCAACCAACATTTGGTTCTTGGCGTGGTGTGCGGGAAGAGAC 1527
QY 1381 TTCACTTCCAAAGCCATCGCCACCGGATTTATCGGGATCAACCCACACTCTCTTCAGGCT 1440
Db 1528 TTCACTTCCAAAGCCATCGCCACCGGATTTATTCGCGATCAACCCACACTCTCTTCAGGCT 1587
QY 1441 CCACCTCGGATGATGAGCAGGACGACATCTTGGATTACTTTGGCAGATGTCAACGTGAAC 1500
Db 1588 CCACCTCGTATGATGAGCAGGAGCAGCATCTTGGATTACTTTGGCAGATGTCAACGTGAAC 1647
QY 1501 AAGCCTCATGTGTGCGCTCCAAAGGATTTGAGCAGCAACCAATCGATAAGCTGCCCAACATC 1560
Db 1648 AAGCCTCATGTGTGCGCTCCAAAGGATTTGAGCAGCTCTATCGATAAGCTGCCCAACATC 1707
QY 1561 AAGGATCTGCCACTGCCACCGGTTCCCGTGACCGCTGAAAGCAGTTGGGCCAGCGCG 1620
Db 1708 AAGGATCTGCCACTGCCACCGGTTCCCGTGACCGCTGAAAGCAGTTGGGCCAGCGCG 1767
QY 1621 TTTGCTCGTGTATCTCCGTGAGCAGGACGCACTGGCAGTTTACTGATACCACTTCCGCGAT 1680
Db 1768 TTTGCTCGTGTATCTCCGTGAGCAGGACGCACTGGCAGTTTACTGATACCACTTCCGCGAT 1827
QY 1681 GCACACGAGCTCTTGTGCGACCGGATCCGCTCATTTGCGACTGAAGCTTGGCGCAGAG 1740
Db 1828 GCACACGAGCTCTTGTGCGACCGGATCCGCTCATTTGCGACTGAAGCTTGGCGCAGAG 1887
QY 1741 GCGGTGCAAAAGCTGACTCTGAGCTTTTGTGCGGAGGCTTGGGGGCGCGCAGCTTAC 1800
Db 1888 GCGGTGCAAAAGCTGACTCTGAGCTTTTGTGCGGAGGCTTGGGGGCGCGCAGCTTAC 1947
QY 1801 GATGTGGCGATGCGCTTTCTCTTTGAGGATTCGTTGGACAGGCTCGAAGCTTGGCGAG 1860
Db 1948 GATGTGGCGATGCGCTTTCTCTTTGAGGATTCGTTGGACAGGCTCGAAGCTTGGCGAG 2007
QY 1861 GCGATGCGCAATGTAAACATTTCAAGATGCTGCTTCGCGGCGCAACACCGTGGGATACACC 1920
Db 2008 GCGATGCGCAATGTAAACATTTCAAGATGCTGCTTCGCGGCGCAACACCGTGGGATACACC 2067
QY 1921 CCGTACCAGACTCCGCTGCTCGCGGCTTTGTTAAGGAAGCTGCGAGCTCGCGGCTGGAC 1980
Db 2068 CCGTACCAGACTCCGCTGCTCGCGGCTTTGTTAAGGAAGCTGCGAGCTCGCGGCTGGAC 2127
QY 1981 ATCTTCCGATCTTTCGACGCGCTTAAAGAGCTCTCCAGATGCGTCCAGCAATCGAGCA 2040
Db 2128 ATCTTCCGATCTTTCGACGCGCTTAAAGAGCTCTCCAGATGCGTCCAGCAATCGAGCA 2187
QY 2041 GTCTGTGAGACCAACACCGCGGTAGCGGCTTATGGCTTATTTCTGTGATCTCTCT 2100
Db 2188 GTCTGTGAGACCAACACCGCGGTAGCGGCTTATGGCTTATTTCTGTGATCTCTCT 2247
QY 2101 GATCCAAATGAAGGCTCTACAACCTGGATTACTACTTAAGATGCGCAGAGGATCGTC 2160
Db 2248 GATCCAAATGAAGGCTCTACAACCTGGATTACTACTTAAGATGCGCAGAGGATCGTC 2307
QY 2161 AAGTCTGGCGCTCAATTTCTGGCCATTAAAGGATATGGCTGGTCTGCTTGGCCAGCTGCG 2220
Db 2308 AAGTCTGGCGCTCAATTTCTGGCCATTAAAGGATATGGCTGGTCTGCTTGGCCAGCTGCG 2367
QY 2221 GTAACAAAGCTGGTCACCGCACTGCGCGGTGAATTCGATCTGCCAGTGCAGTGCACACC 2280
Db 2368 GTAACAAAGCTGGTCACCGCACTGCGCGGTGAATTCGATCTGCCAGTGCAGTGCACACC 2427
QY 2281 CACGACACTGCGGGTGGCCAGTTGGCTTACCTATTTTGTGTCAGCTCAAGCTGGTGCAGAT 2340
Db 2428 CACGACACTGCGGGTGGCCAGTTGGCTTACCTATTTTGTGTCAGCTCAAGCTGGTGCAGAT 2487
QY 2341 GCTGTTGAGCGTGTCTCGCACCACTGTCTGGCACCACTCTCCAGCCATCCCTGTCTGCC 2400
Db 2488 GCTGTTGAGCGTGTCTCGCACCACTGTCTGGCACCACTCTCCAGCCATCCCTGTCTGCC 2547
QY 2401 ATTGTTGCTGCAATTCGCGCACACCGGTGCGGATACCGGTTTGGAGCTCGAGGCTGTTCT 2460

Db	2548	ATTGTTGCTGCATTGCGCACACCGCTGCGGATACCGGTTTGAGCGCTCGAGGCTGTTTCT	2607
Qy	2461	GACCTCGAGCCGTA CTGGGAAGCTGTGCGCGGAGCTGTACTCTGCCATTTGAGTCTGGAACC	2520
Db	2608	GACCTCGAGCCGTA CTGGGAAGCACTGCGCGGAGCTGTACTCTGCCATTTGAGTCTGGAACC	2667
Qy	2521	CCAGGCCCAACCGGTGCGGTCTACCGCCACGAAATCCAGCGGAGACGTTGTCGAACCTG	2580
Db	2668	CCAGGCCCAACCGGTGCGGTCTACCGCCACGAAATCCAGCGGAGACGTTGTCGAACCTG	2727
Qy	2581	CGTGACAGGCCACCGCACTGCGGCTTGCTGATCGCTTTCGAGCTCATCGAAGACAACTAC	2640
Db	2728	CGTGACAGGCCACCGCACTGCGGCTTGCGGATCGTTTCGAACTCATCGAAGACAACTAC	2787
Qy	2641	CGAGCCGTTATGAGATGCTGGGACGCCCAACCAAGGTCAACCCATCTCCAAAGTTGTT	2700
Db	2788	CGAGCCGTTATGAGATGCTGGGACGCCCAACCAAGGTCAACCCATCTCCAAAGTTGTT	2847
Qy	2701	GGCGACCTCGCACTCCACCTGCTGTTGGTGGGGTGTAGATCCAGCAGACTTTGCTGCAGC	2760
Db	2848	GGCGACCTCGCACTCCACCTGCTGTTGGTGGGGTGTAGATCCAGCAGACTTTGCTGCAGC	2907
Qy	2761	CCACAAAAGTACGACATCCGACACTCTGTCTCATCGGTTCTCGCGCGGAGCTTGGTAAC	2820
Db	2908	CCACAAAAGTACGACATCCGACACTCTGTCTCATCGGTTCTCGCGCGGAGCTTGGTAAC	2967
Qy	2821	CCTCAGGTGCTGCCAGAACCACTCGCGACCCCGCGCACTGGAAAGGCGCTCGGAAGGC	2880
Db	2968	CCTCAGGTGCTGCCAGAACCACTCGCGACCCCGCGCACTGGAAAGGCGCTCGGAAGGC	3027
Qy	2881	AAGGCACCTTCGACGGAATTCCTGAGGAAGCAGGCGGCACTCGACGCTGATGATTC	2940
Db	3028	AAGGCACCTTCGACGGAATTCCTGAGGAAGCAGGCGGCACTCGACGCTGATGATTC	3087
Qy	2941	AAGGAACTGCGCAACAGCTCAACCGCTGCTGTTCCGGAAGCCAAACGAAGAGTTCTCTC	3000
Db	3088	AAGGAACTGCGCAACAGCTCAACCGCTGCTGTTCCGGAAGCCAAACGAAGAGTTCTCTC	3147
Qy	3001	GAGCACCTCGCCGCTTCGGAACACCTCTCGCTGGATGATCGTGAATTTCTTACGGA	3060
Db	3148	GAGCACCTCGCCGCTTCGGAACACCTCTCGCTGGATGATCGTGAATTTCTTACGCG	3207
Qy	3061	CTGTCGAGGCGCGGACACTTTGATCCGCTCGCAGATGTCGCGACCCCACTGCTTGT	3120
Db	3208	CTGTCGAGGCGCGGACACTTTGATCCGCTCGCAGATGTCGCGACCCCACTGCTTGT	3267
Qy	3121	CGCTCGATGCGATCTCTGAGCCAGACGATAGGGTATGCGCAATGTTGTGCGCAACGTC	3180
Db	3268	CGCTCGATGCGATCTCTGAGCCAGACGATAGGGTATGCGCAATGTTGTGCGCAACGTC	3327
Qy	3181	AACGCCAGATCCGCCCAATCGGTGCGTGACCGCTCCGTTGAGTCTGTGTCACCGCAACC	3240
Db	3328	AACGCCAGATCCGCCCAATCGGTGCGTGACCGCTCCGTTGAGTCTGTGTCACCGCAACC	3387
Qy	3241	GCAGAAAGGCGAGTTCTTCAACGAGGCGATGTTGCTGCACCAATTCGCTGGTGTGTC	3300
Db	3388	GCAGAAAGGCGAGTTCTTCAACGAGGCGATGTTGCTGCACCAATTCGCTGGTGTGTC	3447
Qy	3301	ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGAGATGCAATGCGCAATCATCGAG	3360
Db	3448	ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGAGATGCAATGCGCAATCATCGAG	3507
Qy	3361	GCTATGAAGATGGAGCAACCAATCACTGCTTCTGTTGACGCGCAAGATTGAAACGGCTGTG	3420
Db	3508	GCTATGAAGATGGAGCAACCAATCACTGCTTCTGTTGACGCGCAAGATTGAAACGGCTGTG	3567
Qy	3421	GTTCTGCTGCAACGAAAGGTGAGGTCGCGACTTGATCGTCGTCGTTTCTTAA	3474
Db	3568	GTTCTGCTGCAACGAAAGGTGAGGTCGCGACTTGATCGTCGTCGTTTCTTAA	3621

RESULT 5
CGPYC

LOCUS	CGPYC	3728 bp	DNA	linear	BCT 08-MAY-1998
DEFINITION	Corynebacterium glutamicum pyc gene.				
ACCESSION	Y09548				
VERSION	Y09548.1				
KEYWORDS	pyc gene; pyruvate carboxylase.				
SOURCE	Corynebacterium glutamicum.				
ORGANISM	Corynebacterium glutamicum.				
REFERENCE	1 (bases 1 to 3728)				
AUTHORS	Peters-Wendisch, P.G., Kreutzer, C., Kalinowski, J., Patek, M., Sahm, H. and Bismann, B.J.				
TITLE	Pyruvate carboxylase from Corynebacterium glutamicum: characterization, expression and inactivation of the pyc gene				
JOURNAL	Microbiology 144 (Pt 4), 915-927 (1998)				
MEDLINE	98240228				
PUBMED	9579065				
REFERENCE	2 (bases 1 to 3728)				
AUTHORS	Peters-Wendisch, P.G.				
TITLE	Direct Submission				
JOURNAL	Submitted (21-NOV-1996) P.G. Peters-Wendisch, Institut fuer Biotechnologie 1, Forschungszentrum Juelich GmbH, Juelich, D-52425, FRG				
FEATURES	Location/Qualifiers				
source	1..3728				
	/organism="Corynebacterium glutamicum"				
	/strain="ATCC 13032"				
gene	109..3587				
	/db_xref="taxon:1718"				
mRNA	109..3587				
	/gene="pyc"				
	/gene="pyc"				
	/evidence="experimental"				
CDS	165..3587				
	/gene="pyc"				
	/BC_number="6.4.1.1"				
	/codon_start=1				
	/transl_table=11				
	/product="pyruvate carboxylase"				
	/protein_id="CAA70739.1"				
	/db_xref="GI:2879823"				
	/db_xref="SPTREMBL:O54587"				
	/translation="MSTHTSSTLPAFKKILVANRGEIIVRAFAALLETGAATVAIIVPR EDSGFHSRFASEAVRIGTSGSPVKAYLDIDELIGAAKKVADAIYPGYFLSENAQL ARECANGITFTGPTPEVLDLTDGKSRVTAARKAGLPVLAESTPSKNIDELVKSAG QTPYFVKAVAGGGRMRFVASPELRKLATSEAREAAFGDGVYVERAVINPOH IEVQIIGDHTGEVHLRYERDCSLORRHQKVEIAPAQHLDPDLDRICADAVKFSI GQAGTVEFLVDEKHNHVFIEBNPRIQVEHTVEVTEVDLVAQMRLAAGATLKEI GLTQDKIKTHGAALQCRITTEDPNNRFRPDITITATSPFGAGVLRDDGAAQGGBIT AHFDSMLVMTCRGSDFTAVARAQALAEFTVSGVATNIGFLRLAREEDFTSKRIA TGFIADPHLLQAPPADDEQRIIDYLDVTVNKHGVRPKDVAAPIDKLPNTKDLPL PRGSRLLKQLGPAAPARDLREODALAVDTTFRDAHQSLLATRVRSFALKPAAEAVA KLTPELLSEAMGGATYDVAMRFLFEDPWRDLDELREAMPNVIQMLRGRNTVGTP YDSCVRAFKERASSGVDFRIFDALNDVSQMRPAIDAVLENTAVAEVMAVMSGL SPNKKLYLDLYLWAEIIVSGAHILAKDMAGLLRPAATKLVTLURREFDLPVH VHTHDGQGLATYFAAQAGADAVDGAAPLSGTSQPSLSAIVAAFAHTRDRLGLS LEAVSLEPYNEAVRGLYLPFFSGTGPGRVYRHEIPGQSLNRAQATLGLADFL LEEDNVAVNEMLGRPTKTPSSKVVGDLALHVGADPADPADPKDYDIPDSVI AFRLGRLNPGCGWPEPLTRALEGSECKAPLVEPSEQAHLDADDSKERENSLNR LPLFPTEEFLEHRRRFGNTSALDDREFYGLVEGRETLIRLPDVTPLLVRLDLASE PDDKMRNVANVNGQIRPMRVDRSVSTATAEKADSNKHGVAAPFAGVVTIVA EGDEVKAGDAVAIIIEAMKMEATITASVDGKIDRVVVVPAATKVEGGDLIVVVS"				
BASE COUNT	813 a	1060 c	1034 g	821 t	
ORIGIN					

Query Match 97.8%; Score 3398.8; DB 1; Length 3728;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3427; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy	1	GTGACTGCTATCACCCCTTGGCGGTCTCTCTGTGAAGGAATAATTACTCTAGTCTGCACCT	60
Db	114	GTGACTGCTATCACCCCTTGGCGGTCTCTCTGTGAAGGAATAATTACTCTAGTCTGCACCT	173

Db 2334 GTAACCAAGCTGGTCAACCGCACTGCGCGTGAAATTCGATCTGCCAGTGCACGTGCACACC 2333
Qy 2281 CACGACACTGCGGGTGGCCAGTTGGCTACTCTAATTTGCTGCGAGCTCAAGCTGGTGCAGAT 2340
Db 2394 CACGACACTGCGGGTGGCCAGCTGGCAACCTACTATTGCTGCGAGCTCAAGCTGGTGCAGAT 2453
Qy 2341 GCTGTTGACGGTGCTTCCGCAACCACTGTCGGCACCACCTCCAGCCATCCTGTCGACC 2400
Db 2454 GCTGTTGACGGTGCTTCCGCAACCACTGTCGGCACCACCTCCAGCCATCCTGTCGACC 2513
Qy 2401 ATTGTTGCTGCTATTCGCGCACACCGCTCGCGATACCGGTTTGAGCTCGAGCTGTTTCT 2460
Db 2514 ATTGTTGCTGCTATTCGCGCACACCGCTCGCGATACCGGTTTGAGCTCGAGCTGTTTCT 2573
Qy 2461 GACCTCGAGCCGTAATCGGAAGCTGTGCGCGGACTGTACTCTGCCATTTGAGCTTGGAAACC 2520
Db 2574 GACCTCGAGCCGTAATCGGAAGCACTGCGCGGACTGTACTCTGCCATTTGAGCTTGGAAACC 2633
Qy 2521 CCAGGCCCAACCGGTGCGTCTACCGCCACGAAATCCAGGCGGACAGTTGTCCAACTG 2580
Db 2634 CCAGGCCCAACCGGTGCGGCTTACCGCCACGAAATCCAGGCGGACAGTTGTCCAACTG 2693
Qy 2581 CGTGCAAGGCCACCGCACTGGGCTTCTGATCGCTTCGAGCTCATCGAAGACAACCTAC 2640
Db 2694 CGTGCAAGGCCACCGCACTGGGCTTCTGGAATCTTCGAACTCATCGAAGACAACCTAC 2753
Qy 2641 CGACCGTTAATGAGATCTGGGACGCCCAACCAAGGTCAACCCATCTCCAAAGGTTGTT 2700
Db 2754 GCAGCCGTTAATGAGATCTGGGACGCCCAACCAAGGTCAACCCATCTCCAAAGGTTGTT 2813
Qy 2701 GCGACCTCGCACTCCACTGGTTGGTGCGGGTGTAGATTCAGCAGAGACTTGTCTGCAGAC 2760
Db 2814 GCGACCTCGCACTCCACTGGTTGGTGCGGGTGTGATTCAGCAGAGACTTGTCTGCAGAT 2873
Qy 2761 CCACAAAAGTACGATCCCAAGACTCTGTCATCGCTTCGCTCGCGGGGAGCTGGTAAAC 2820
Db 2874 CCACAAAAGTACGATCCCAAGACTCTGTCATCGCTTCGCTCGCGGGGAGCTGGTAAAC 2933
Qy 2821 CCTCCAGTGGCTGGCCAGAACCACTGGCGACCCCGCGCACTCGAAGGCGCTCCGAAGGC 2880
Db 2934 CCTCCAGTGGCTGGCCAGAGCACTGGCGACCCCGCGCACTCGAAGGCGCTCCGAAGGC 2993
Qy 2881 AAGGCACCTTCGACGGAGTTCTTGAGGAAGACGAGGGCGACCTCGAGCGCTGATGATTC 2940
Db 2994 AAGGCACCTTCGACGGAGTTCTTGAGGAAGACGAGGGCGACCTCGAGCGCTGATGATTC 3053
Qy 2941 AAGGAACTGCGCAACAGCCTCAACCGCTGCTGTTCCGGAAGCCCAACGGAAGGTTCTC 3000
Db 3054 AAGGAACTGCGCAATAGCCTCAACCGCTGCTGTTCCGGAAGCCCAACGGAAGGTTCTC 3113
Qy 3001 GAGCACCTGCGCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTTCTTACGGA 3060
Db 3114 GAGCACCTGCGCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTTCTTACGCG 3173
Qy 3061 CTGCTGAGGGCCGGGAGACTTTGATCGCTTCGCGAGATGTGGCACCCTACTGCTGTT 3120
Db 3174 CTGCTGAGGGCCGGGAGACTTTGATCGCTTCGCGAGATGTGGCACCCTACTGCTGTT 3233
Qy 3121 CGCTTGGATGCTCTGAGCCAGACGATGAAGGTAAGCGCAATGTTGTGGCCCAACGTC 3180
Db 3234 CGCTTGGATGCTCTGAGCCAGACGATGAAGGTAAGCGCAATGTTGTGGCCCAACGTC 3293
Qy 3181 AACGGCCAGATCCCGCCCAATCGTGTGCGTGACCGCTCCGTTGAGTCTGTCAACCGCAACC 3240
Db 3294 AACGGCCAGATCCCGCCCAATCGTGTGCGTGACCGCTCCGTTGAGTCTGTCAACCGCAACC 3353
Qy 3241 GCAGAAAAGGAGATTTCTTCCAAAGGGCCATGTTGCTGCAACCATTCGCTGGTGTGTC 3300
Db 3354 GCAGAAAAGGAGATTTCTTCCAAAGGGCCATGTTGCTGCAACCATTCGCTGGTGTGTC 3413
Qy 3301 ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGAGATGCAAGTGGCAATCATCGAG 3360
Db 3414 ACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGAGATGCAAGTGGCAATCATCGAG 3473

Qy 3361 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGCAAGATTTGAACGGTTGTG 3420
Db 3474 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGCAAAATCGATCGGTTGTG 3533
Qy 3421 GTTCCTCTCTCAACGAAGGTGGAAGGTGGCAGCTTGATCGTCGTCGTTCTCTAA 3474
Db 3534 GTTCCTCTCTCAACGAAGGTGGAAGGTGGCAGCTTGATCGTCGTCGTTCTCTAA 3587
RESULT 6
LOCUS A97276 Sequence 1 from Patent WO9918228. 3728 bp DNA linear PAT 26-JAN-2000
DEFINITION A97276
ACCESSION A97276
VERSION A97276.1 GI:6780661
KEYWORDS
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 3728)
AUTHORS Peters-Wendisch, P. and Eikmanns, B.
TITLE METHOD FOR MICROBIAL PRODUCTION OF AMINO ACIDS OF THE ASPARTATE
AND/OR GLUTAMATE FAMILY AND AGENTS WHICH CAN BE USED IN SAID METHOD
JOURNAL Patent: WO 9918228-A 1 15-APR-1999;
PETERS WENDISCH PETRA (DE); EIKMANNs BERND (DE)
FEATURES
Location/Qualifiers
source 1..3728
/organism="unclassified"
/db_xref="taxon:32844"
BASE COUNT 813 a 1060 c 1034 g 821 t
ORIGIN
Query Match 97.8%; Score 3398.8; DB 6; Length 3728;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3427; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
Qy 1 GTGACTGCTATCACCCCTTGGCGGTCTCTGTTGAAAGAAATAATTACTCTAGTGTGCACT 60
Db 114 GTGACTGCTATCACCCCTTGGCGGTCTCTGTTGAAAGAAATAATTACTCTAGTGTGCACT 173
Qy 61 CACACATCTTCAACGGTTCAGCATTTCAAAAAGATCTTGGTAGCAAAACCGGGGAAAATC 120
Db 174 CACACATCTTCAACGGTTCAGCATTTCAAAAAGATCTTGGTAGCAAAACCGGGGAAAATC 233
Qy 121 GCGTCCGTCGTTTCCGTGCGACACTCGAAACCGGTGACGCGTAGCTATTTCACCC 180
Db 234 GCGTCCGTCGTTTCCGTGCGACACTCGAAACCGGTGACGCGTAGCTATTTCACCC 293
Qy 181 CGTGAAGATCGGGGATCATTTCCACCGCTCTTTTGTCTTCTGAAGCTGTCGCAATTGGTACT 240
Db 294 CGTGAAGATCGGGGATCATTTCCACCGCTCTTTTGTCTTCTGAAGCTGTCGCAATTGGTACT 353
Qy 241 GAAGGCTCACAGTCAAGGCGTACCTTGGACATCGATGAAATTTATCGGTGAGCTTAAAAA 300
Db 354 GAAGGCTCACAGTCAAGGCGTACCTTGGACATCGATGAAATTTATCGGTGAGCTTAAAAA 413
Qy 301 GTTAAAGCAGATGCTATTTTACC CGGGATATGGCTTCTGCTGAAATGCCAGCTTGGC 360
Db 414 GTTAAAGCAGATGCTATTTTACC CGGGATATGGCTTCTGCTGAAATGCCAGCTTGGC 473
Qy 361 CGCGAGTCCGCGGAAAACGGCATTTACTTTTATTTGGCCCAACCCAGAGGTTCTTTGATCTC 420
Db 474 CGCGAGTCCGCGGAAAACGGCATTTACTTTTATTTGGCCCAACCCAGAGGTTCTTTGATCTC 533
Qy 421 ACCGGTGATTAAGTCTCTGTCGGGTAAACCGCGCAAGAGGCTGGTCTGCCAGTTTGGCG 480
Db 534 ACCGGTGATTAAGTCTCTGTCGGGTAAACCGCGCGAAGAGGCTGGTCTGCCAGTTTGGCG 593
Qy 481 GAATCCACCCCGAGCAAAAACATCGATGACATCGTTTAAAGCGCTGAAGCCAGACTTAC 540
Db 594 GAATCCACCCCGAGCAAAAACATCGATGACATCGTTTAAAGCGCTGAAGCCAGACTTAC 653

Qy	541	CCATCTTTGTAAGGCACGTTCGCCGGTGGCGGACGCGGTATGCGCCTTTGTTCTTCA	600
Db	654	CCCATCTTTGTGAAGGCAGTTGCCGGTGGCGGACGCGGTATGCGCTTTTGTTGCTTCA	713
Qy	601	CCTGATGAGCTCCGCAAAATTGGCAACAAGAACATCTCGTGAAGCTGAAGCGCATTCGGC	660
Db	714	CCTGATGAGCTTCGCAAAATTAGCAACAAGAAGCATCTCGTGAAGCTGAAGCGCGCTTCGGC	773
Qy	661	GACGGTTCCGTTATGTGCAACGCTGCTGATTAAACCCCGCACATTTGAAGTGCAGATC	720
Db	774	GATGGCGGGTATATGTGCAACGCTGCTGANTAAACCCCTCAGCATATTTGAAGTGCAGATC	833
Qy	721	CTTGGCGATCGCACTGGAGAAAGTTGTATACACTTTTATGAACGTGACTCTCACTCAAGCGT	780
Db	834	CTTGGCGATCACACTGGAGAAAGTTGTATACACTTTTATGAACGTGACTCTCACTCAAGCGT	893
Qy	781	CGTCAACCAAAAGTTGTCGAANAATTGGCCAGACACAGCATTTTGGATCCAGRACTCGGTGAT	840
Db	894	CGTCAACCAAAAGTTGTCGAANAATTGGCCAGACACAGCATTTTGGATCCAGRACTCGGTGAT	953
Qy	841	CGCANTTGCGGGATGAGTAAAGTTCTGCGCGCTCCATGGTTTACAGGGCGCGGGAAACC	900
Db	954	CGCANTTGCGGGATGAGTAAAGTTCTGCGCGCTCCATGGTTTACAGGGCGCGGGAAACC	1013
Qy	901	GTGGAAATCTTGCTGATGAAGAGGGCAACACGTTTTTTCATCGAANAATGAACCCAGTATC	960
Db	1014	GTGGAAATCTTGCTGATGAAGAGGGCAACACGTTTTTTCATCGAANAATGAACCCAGTATC	1073
Qy	961	CAGGTTGAGCACACCGTGACTGAAGAAAGTCAACGAGGTGGACCTGGTGAAGCGCCAGATG	1020
Db	1074	CAGGTTGAGCACACCGTGACTGAAGAAAGTCAACGAGGTGGACCTGGTGAAGCGCCAGATG	1133
Qy	1021	CGCTTGCGTCTGGTGCAAACCTTGAAGGAAATCGGCTCTGACCCAAGATAAGATCAAGACC	1080
Db	1134	CGCTTGCGTCTGGTGCAAACCTTGAAGGAAATCGGCTCTGACCCAAGATAAGATCAAGACC	1193
Qy	1081	CACGSGTGACGACTGCAGTGGCGGATCAACAACGGAAGATCCAAAACACCGGCTTCGCGCCA	1140
Db	1194	CACGSGTGACGACTGCAGTGGCGGATCAACAACGGAAGATCCAAAACACCGGCTTCGCGCCA	1253
Qy	1141	GATACCGGAACTATCACCGCGTACCGCTCAACGAGGCGGAGCTGGGCTTCGTTGACGGT	1200
Db	1254	GATACCGGAACTATCACCGCGTACCGCTCAACGAGGCGGAGCTGGGCTTCGTTGACGGT	1313
Qy	1201	GCAGCTCAGCTCGGTGGCGAAATCAACCGCACACTTTTGACTCCATGCTGTAAGAAATGACC	1260
Db	1314	GCAGCTCAGCTCGGTGGCGAAATCAACCGCACACTTTTGACTCCATGCTGTAAGAAATGACC	1373
Qy	1261	TGCGGNGHTTCGACCTTTGAAACTGCTGTGCTGCGACAGCGCGCTTGCTGAGTTTC	1320
Db	1374	TGCGGNGHTTCGACCTTTGAAACTGCTGTGCTGCGACAGCGCGCTTGCTGAGTTTC	1433
Qy	1321	ACCGTCTCGGTGTGAACCAACATTTGGTTCTTCGGTGGCTTCTCGCGGAGAGGAC	1380
Db	1434	ACCGTCTCGGTGTGAACCAACATTTGGTTCTTCGGTGGCTTCTCGCGGAGAGGAC	1493
Qy	1381	TTCACTTCOAGCGCATCGCACCGGATTTATCGGGATCAACGCACTCTCTTCAGGCT	1440
Db	1494	TTCACTTCOAGCGCATCGCACCGGATTTATTCGGCATCACCGCATCTCTTCAGGCT	1553
Qy	1441	CCACTGCGGATGATGAGAGGAGCGCATCTCTGATTTACTTTGGCAGATGTCAACCGTAGAC	1500
Db	1554	CCACTGCGGATGATGAGAGGAGCGCATCTCTGATTTACTTTGGCAGATGTCAACCGTAGAC	1613
Qy	1501	AAGCCTCATGGTTCGCTCCAAAGATGTTGCGACCAACCAATCGATAAGCTGCCCAAATC	1560
Db	1614	AAGCCTCATGGTTCGCTCCAAAGATGTTGCGAGCTCTTATCGATAAGCTGCCCAAATC	1673
Qy	1561	AAGGATCTGCCACTGCCACGCGGTTCCGCTGAACCGCTTGAAGCAGCTTGGCCACGCGCG	1620
Db	1674	AAGGATCTGCCACTGCCACGCGGTTCCGCTGAACCGCTTGAAGCAGCTTGGCCACGCGCG	1733
Qy	1621	TTTGTCTGCTGATCTCCGTGAGCAGGACGCGACTGCGAGTCTACTGATACACCTTCCGCGAT	1680


```
gene
CDS
4715..5629
/gene="Cg10658"
4715..5629
/gene="Cg10658"
/codon_start=1
/transl_table=11
/product="PEP phosphonotomutase and related enzymes"
/protein_id="BAB98051.1"
/db_xref="GI:21323424"
/translation="MSSAVAPTERKALRALAAPEIARMPGAFSPLAARAOEAGF
EGVTVGAVAADIALDGLITLTITEVAHRSQIARVLDVLVDATGTGGEPMGAR
RDSLEAGVAGHLEQDNRKRGHLDGKGVGIDINVRRIARAVERDEQVICA
RTDAGVEGDSALERAKYADACAGDMIFTEALYIPDFEKPRAVDIPLIANITPFG
YELLPAQLEDIGYNAYIYPTLLRIAMGVEQALGDIAITGIQTVDWRMQHRSKL
YELRYNEYNAFDQQVFTYSADSKPIF"
5658..6809
/gene="Cg10659"
5658..6809
/gene="Cg10659"
/EC_number="4.1.1.37"
/notes="PF00285:Citrate synthase"
/codon_start=1
/transl_table=11
/product="Citrate synthase"
/protein_id="BAB98052.1"
/db_xref="GI:21323425"
/translation="MSSATTDTKGLYGIADYTAVKVMPETNSLTIRGYAVEDIV
ENCSEFVGTDLVHGEPLTQALAEFNERGRSVRLDAGLSLIHSLPKAHPMQVNR
TAVSYMGTQSEYFTTDSHRIKRVGHTLLAQLPMVLAMDIRRKGLDIIAPDSKVA
ENGLSMVGTGTPSPASNPADVRDFEKSILVAEHSFNASTFTARVITSTKSDVYSAT
TGALGKGLPHGSGANFVMTWLAIDDPNKAAMINNALDNKNVWFGFHRVYRGD
SRVSPMEKSFRELAARDHGVKWMYENMRDAMDARTGICKNLOPPAGPAVHLGFPV
DFTPLFVIARVAGTAAHIVEQYENNSLRPLSEYNGEVEQVAPIER"
complement(6846..7040)
/gene="Cg10660"
complement(6846..7040)
/gene="Cg10660"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB98053.1"
/db_xref="GI:21323426"
/translation="MYFDSQDVDDLHFSFPELVADQALDAHRCAPHFYVWREAVA
KHVLPQSQKNVGTFRMLHHG"
complement(7129..8172)
/gene="Cg10661"
complement(7129..8172)
/gene="Cg10661"
/EC_number="1.1.1.82"
/notes="PF02615:Malate/L-lactate dehydrogenase"
/codon_start=1
/transl_table=11
/product="Malate/L-lactate dehydrogenases"
/protein_id="BAB98054.1"
/db_xref="GI:21323427"
/translation="MPSEVTYNAQQLTVLCTDILTITKGVPAADHLVGDLSLVQADLWGH
PSHGVLRLPWVRLHSGAMTHAHVEVLDLGLVLDGHNIGQVLADHARKEAVT
RAMFGIGAVSRNSHFGTAMVYTRKAAQGCVSILTTNASPAMAPWGGRKRGIN
PKSLAAPGETAVTDIANTAVARGKIYHARTNNMIPETWALTSEGATPTDPAEIN
GVWLPMAHGKYALSPMWDVLSGLVGSOSHSTKVHGPYDPTPPGAGHLFTALDVAF
RDQDDPDAISDLAGEVKSIPKAQNTTEEIFYGESEDRHAKNSHGISLPEKTMWEL
QELAIENHVTH"
complement(8255..8987)
/gene="Cg10662"
complement(8255..8987)
/gene="Cg10662"
/notes="PF00392:Bacterial regulatory proteins, gntR family"
/codon_start=1
/transl_table=11
/product="Transcriptional regulators"
/protein_id="BAB98055.1"
/db_xref="GI:21323428"
/translation="MVQKDAQSPATRKADQVYVYQIRRIEDGTINPCQRMSEVWLVE
```

```
HTGASRTVPRDALRELADELIIILEPROQPMVSPLSLRHIKDLFEFERRIVEVAALBEI
SVGASKSPRIFGEFTSLAADFRELENSADDADFTADFRRLTSKFDLDVAANTHNOFLG
RSILSLKPTHTRLRIIAHSDHARLRQSVQEHCEAVASGDLRSAGACROHLIHVE
KSILTALINADSTGSGQIDIRS"
complement(9094..9798)
/gene="Cg10663"
complement(9094..9798)
/gene="Cg10663"
/codon_start=1
/transl_table=11
/product="Hypothetical membrane protein"
/protein_id="BAB98056.1"
/db_xref="GI:21323429"
/translation="MRFSRVLPALLITITAVSIPPTASAATLTADTKELCIASNTDSEA
VTFWNSIDSVREQLDELQDQPGIKAAESITADQDNAPPAEQLQVRLAIEAGE
GLAWLPDDPTILADPNABESFKIETIYDEAKDIIISGFSSDPASVLSQLQAQATGTR
TABIRAEVADRTDDDDNESQTALKEDFQNCIDAIDARPIPLQIILIGGALAIVL
GIRAWTNSRKQSKHSQ"
complement(9878..12094)
/gene="Cg10664"
complement(9878..12094)
/gene="Cg10664"
/EC_number="1.1.1.42"
/notes="TIGR00178:monomer_idh: isocitrate dehydrogenase,
NADP-dependent"
/codon_start=1
/transl_table=11
/product="Monomeric isocitrate dehydrogenase (NADP+)"
/protein_id="BAB98057.1"
/db_xref="GI:21323430"
/translation="NAKIIWTRTDRAPLLATYSLKPVVEAFATAGIEVETRIDISLAG
RIIAQPERLTDDQKVALAEGLAKELATPEBANIILKPNISASVPOLKAAIKELQDOG
YDIPEDNATDDEKDIILARVAKGSAPVNLREGNSDRRAPIAKVNFKVFPFHRM
GENSADKSTNATVMTDANDFRHNEKSIILDADEADEVQIKHIAADGTETILKDSLLLEGE
```

Query Match	97.8%	Score 3398.8	DB 1	Length 332050
Best Local Similarity	98.6%	Pred. No. 0		
Matches 3427	Conservative	Mismatches 47	Indels 0	Gaps 0
QY 1	GTGACTGTATCACCTTGGGGTCTCTTGTGAAGGAATTAATCTCTAGTGTGCACT 60			
DB 37210	GTGACTGTATCACCTTGGGGTCTCTTGTGAAGGAATTAATCTCTAGTGTGCACT 37269			
QY 61	CACACATCTCAAGCTTCAGCATTCAAAAGATCTTGGTAGCAACCGCGGCAATC 120			
DB 37270	CACACATCTCAAGCTTCAGCATTCAAAAGATCTTGGTAGCAACCGCGGCAATC 37329			
QY 121	GGGTCCTGTCTTTCGTTGACGACCTCGAAAACCGGTGACGACCGGTAGCTATTTACCCC 180			
DB 37330	GGGTCCTGTCTTTCGTTGACGACCTCGAAAACCGGTGACGACCGGTAGCTATTTACCCC 37389			
QY 181	CGTGAAGATCGGGGATCATTCACCGCTCTTTTGTCTTCTGAAGCTGTCCGCAATGGTACT 240			
DB 37390	CGTGAAGATCGGGGATCATTCACCGCTCTTTTGTCTTCTGAAGCTGTCCGCAATGGTACT 37449			
QY 241	GAAGGCTCACAGTCAAGGGGTACTCGACATCGATGAAATATCGGTGACGCTAAAAA 300			
DB 37450	GAAGGCTCACAGTCAAGGGGTACTCGACATCGATGAAATATCGGTGACGCTAAAAA 37509			
QY 301	GTTAAAGCAGATGCTATTATCCCGGGATATGGCTTCTCTGTGAAAAATGCCAGCTTCCC 360			
DB 37510	GTTAAAGCAGATGCTATTATCCCGGGATATGGCTTCTCTGTGAAAAATGCCAGCTTCCC 37569			
QY 361	CGCGAGTCGGGGAAAACGGCATTTCTTTTATGGCCCAACCCAGAGTCTTCTGATCTC 420			
DB 37570	CGCGAGTCGGGGAAAACGGCATTTCTTTTATGGCCCAACCCAGAGTCTTCTGATCTC 37629			
QY 421	ACCGGTGATAAGTCTGTGCGGTAAACCGCGCGGAAAGCGTGGTCTGCCAGTTTGGCG 480			
DB 37630	ACCGGTGATAAGTCTGTGCGGTAAACCGCGCGGAAAGCGTGGTCTGCCAGTTTGGCG 37689			
QY 481	GAATCCACCCGAGCAAAAACATCGATGACATCGTTAAAGCGCTGAGGCCAGACTTAC 540			

Db 37690 GAATCCACCCGAGCAAAAAATCGATGAGATCGTTAAAGCGCTGAAGGCCAGACTTAC 37749
Qy 541 CCCATCTTTGTAAGCGAGTTGCCGGTGGTGGCGGACCGGTATGCGCTTTGTTCTTCA 600
Db 37750 CCCATCTTTGTAAGCGAGTTGCCGGTGGTGGCGGACCGGTATGCGCTTTGTTGCTTCA 37809
Qy 601 CCTGATGAGCTCCGCNAATTTGGCAACAGAAAGCATCTCGTGAAGCTGAAGCGGATTCGGC 660
Db 37810 CCTGATGAGCTTCGCAATTAAGCAACAGAAAGCATCTCGTGAAGCTGAAGCGGCTTCGGC 37869
Qy 661 GACGGTTCCGATATGTCGAACGTCGTGTATTAACCCAGCAATTTGAAGTGCAGATC 720
Db 37870 GATGCGCGGTATATGTCGAACGTCGTGTATTAACCCAGCAATTTGAAGTGCAGATC 37929
Qy 721 CTTGGCGATGCCATGGAGAAATTTGTACACCTTTATGAACGTGATCTCACTCAGCGT 780
Db 37930 CTTGGCGATCACCTGGAGAAATTTGTACACCTTTATGAACGTGATCTCACTCAGCGT 37989
Qy 781 CGTCACCAAAAGTTGCGAAATTTGCCGACAGCAGCATTTGGATCCAGAACTCGGTGAT 840
Db 37990 CGTCACCAAAAGTTGCGAAATTTGCCGACAGCAGCATTTGGATCCAGAACTCGGTGAT 38049
Qy 841 CGCATTTGTGCGGATGAGTAAGTTCTGCGCTCCATTGGTTACAGGGCGCGGGAACC 900
Db 38050 CGCATTTGTGCGGATGAGTAAGTTCTGCGCTCCATTGGTTACAGGGCGCGGGAACC 38109
Qy 901 GTGAAATTTGTGCGATGAAAGGGCAACACAGTTTTCATCGAAATGAACCCACGTATC 960
Db 38110 GTGAAATTTGTGCGATGAAAGGGCAACACAGTTTTCATCGAAATGAACCCACGTATC 38169
Qy 961 CAGGTTGAGCACCGGTGACTGAAGAGTCAACGAGTGGACCTGGTGAAGGCGGAGATG 1020
Db 38170 CAGGTTGAGCACACCGGTGACTGAAGAGTCAACGAGTGGACCTGGTGAAGGCGGAGATG 38229
Qy 1021 CGTTTGGCTGCTGTGCAACCTTGAAGAAATTTGGTCTGACCCAGATTAAGATCAAGACC 1080
Db 38230 CGTTTGGCTGCTGTGCAACCTTGAAGAAATTTGGTCTGACCCAGATTAAGATCAAGACC 38289
Qy 1081 CACGGTCAGCACTGCAGTGCAGATCACCAAGGAATCCAAACACCGGCTTCGGCCCA 1140
Db 38290 CACGGTCAGCACTGCAGTGCAGATCACCAAGGAATCCAAACACCGGCTTCGGCCCA 38349
Qy 1141 GATACCGAACTATCACCGGTACCGTCAACAGCGGAGCTGGCGTTGCTTGAAGGT 1200
Db 38350 GATACCGAACTATCACCGGTACCGTCAACAGCGGAGCTGGCGTTGCTTGAAGGT 38409
Qy 1201 GCAGTCAGCTCGGTGGGGAATCACCGCACACTTTGACTCCATGCTGGTGAABATGACC 1260
Db 38410 GCAGTCAGCTCGGTGGGGAATCACCGCACACTTTGACTCCATGCTGGTGAABATGACC 38469
Qy 1261 TGCCGTGGTTCGAGCTTTGAAACTGCTGTTGCTCGTGCACACGCGGCTTGGCTGAGTTC 1320
Db 38470 TGCCGTGGTTCGAGCTTTGAAACTGCTGTTGCTGTCGACAGCGCGGCTTGGCTGAGTTC 38529
Qy 1321 ACCGTGCTGGTGTGCAACCAATTTGGTTCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1380
Db 38530 ACCGTGCTGGTGTGCAACCAATTTGGTTCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGT 38589
Qy 1381 TTCACTTCCAGCGCATCGCCACCGGATTTATCGCGATCACCCACACCTCTTTCAGGCT 1440
Db 38590 TTCACTTCCAGCGCATCGCCACCGGATTTATCGCGATCACCCACACCTCTTTCAGGCT 38649
Qy 1441 CCACCTCGGATGATGACGAGGACCGATCTCGATTTACTTTGGCAGATGTCAACCGTGAAC 1500
Db 38650 CCACCTCGGATGATGACGAGGACCGATCTCGATTTACTTTGGCAGATGTCAACCGTGAAC 38709
Qy 1501 AAGCCTCATGGTGTGCTGCAAGAGATTTGACAGCAACCAATTCGATAAGCTGCCCAATC 1560
Db 38710 AAGCCTCATGGTGTGCTGCAAGAGATTTGACAGTCTCTATCGATAAGCTGCCCAATC 38769
Qy 1561 AAGGATCTGCCATGCCACCGGTTCCGCTGACCGCTGAAGAGCTTGGCCGAGCGCG 1620
Db 38770 AAGGATCTGCCATGCCACCGGTTCCGCTGACCGCTGAAGAGCTTGGCCGAGCGCG 38829

Qy 1621 TTTGCTCGTGATCTCCGTGAGCAGGACGCACTGSCAGTTACTGTATACACCTTCGCGAT 1680
Db 38830 TTTGCTCGTGATCTCCGTGAGCAGGACGCACTGSCAGTTACTGTATACACCTTCGCGAT 38889
Qy 1681 GCACACAGCTTTGCTTGGACCCGAGTCCGCTCATTTCCGACTGAAGCCTTCGCGCAG 1740
Db 38890 GCACACAGCTTTGCTTGGACCCGAGTCCGCTCATTTCCGACTGAAGCCTTCGCGCAG 38949
Qy 1741 GCCCTCCAAAGCTGACTCCTGAGCTTTTGTCTCGTGAAGGCTTGGGCGCGCACCTAC 1800
Db 38950 GCCCTCCAAAGCTGACTCCTGAGCTTTTGTCTCGTGAAGGCTTGGGCGCGCACCTAC 39009
Qy 1801 GATGTGCGGATGCGTTTCTCTTTGAGGATCCGTFGGACAGGCTCGACGAGCTGCGGAG 1860
Db 39010 GATGTGCGGATGCGTTTCTCTTTGAGGATCCGTFGGACAGGCTCGACGAGCTGCGGAG 39069
Qy 1861 GCGATGCCGAATGTAAACATTTAGATGCTGCTTCGCGCGCAACACCGTGGGATACACC 1920
Db 39070 GCGATGCCGAATGTAAACATTTAGATGCTGCTTCGCGCGCAACACCGTGGGATACACC 39129
Qy 1921 CCCTACCCAGACTCCGTCTCGCGCGCTTTTAAAGAAAGCTGCAGCTCCGCGGTGAC 1980
Db 39130 CCCTACCCAGACTCCGTCTCGCGCGCTTTTAAAGAAAGCTGCAGCTCCGCGGTGAC 39189
Qy 1981 ATCTTCCGCACTTTGACGCGCTTAAACGAGCTCTCCAGATGCGTCCAGCAATCGACCA 2040
Db 39190 ATCTTCCGCACTTTGACGCGCTTAAACGAGCTCTCCAGATGCGTCCAGCAATCGACCA 39249
Qy 2041 GTCTTGAGAGCAACACCGCGGTAGCCGAGTGGCTATGCTTATTTCTGTGATCTCTCT 2100
Db 39250 GTCTTGAGAGCAACACCGCGGTAGCCGAGTGGCTATGCTTATTTCTGTGATCTCTCT 39309
Qy 2101 GATCCAAATGAAGCTCTACACCTCGATTACTACCTAAAGATGGCAGAGAGATCGTC 2160
Db 39310 GATCCAAATGAAGCTCTACACCTCGATTACTACCTAAAGATGGCAGAGAGATCGTC 39369
Qy 2161 AAGTCTGGCTCACTTTCTGGCCATTAAGGATATGCTTGTCTTGTCCGCCAGCTGCG 2220
Db 39370 AAGTCTGGCTCACTTTCTGGCCATTAAGGATATGCTTGTCTTGTCCGCCAGCTGCG 39429
Qy 2221 GTAACCAAGCTGTCAACGCACTGCGCGCTGMAATTTCATCTGCCAGTGCAGCTGCACACC 2280
Db 39430 GTAACCAAGCTGTCAACGCACTGCGCGCTGMAATTTCATCTGCCAGTGCAGCTGCACACC 39489
Qy 2281 CACGACACTCGGCTGCGCAGTTGGCTACCTTCTGCTGAGCTCAAGCTTGGTGCAGAT 2340
Db 39490 CACGACACTCGGCTGCGCAGTTGGCTACCTTCTGCTGAGCTCAAGCTTGGTGCAGAT 39549
Qy 2341 GCTGTTGACGGTCTTCGCGACCACTGCTTGGCAACCACTCCAGCCATCCCTGCTGCC 2400
Db 39550 GCTGTTGACGGTCTTCGCGACCACTGCTTGGCAACCACTCCAGCCATCCCTGCTGCC 39609
Qy 2401 ATTGTTGCTGATTCGCGCACACCGCTCGGATACCGGTTTGAAGCTTGAAGCTTGTCT 2460
Db 39610 ATTGTTGCTGATTCGCGCACACCGCTCGGATACCGGTTTGAAGCTTGAAGCTTGTCT 39669
Qy 2461 GACCTCGAGCCGCTACTCGGGAAGCTGTCGCGGACTGTACCTGCCATTTGAGTCTGGAACC 2520
Db 39670 GACCTCGAGCCGCTACTCGGGAAGCTGTCGCGGACTGTACCTGCCATTTGAGTCTGGAACC 39729
Qy 2521 CCAGGCCCAACCGGCTCGCTTACCGCCAGCAAAATCCAGCGGACAGTGTGTCCAAACCTG 2580
Db 39730 CCAGGCCCAACCGGCTCGCTTACCGCCAGCAAAATCCAGCGGACAGTGTGTCCAAACCTG 39789
Qy 2581 CGTGCAAGCGCACCGCACTGGGCTTGTGATCGCTTTCAGCTTCATCGAAGCAACTAC 2640
Db 39790 CGTGCAAGCGCACCGCACTGGGCTTGTGATCGCTTTCAGCTTCATCGAAGCAACTAC 39849
Qy 2641 GCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCATCTCCAGGTTGTT 2700
Db 39850 GCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCATCTCCAGGTTGTT 39909

Db 106060 GTGGAAATTCCTGGTCGATGAAGAGGGCAACACCGTCTTCATCGAAATGAACCCACGGTATC 106119
Qy 961 CAGGTTGAGCACACCGTGACTGAAGAAGTCAACCGAGGTGGACCTGGTGAAGCGCAGATG 1020
Db 106120 CAGGTTGAGCACACCGTGACTGAAGAAGTCAACCGAGGTGGACCTGGTGAAGCGCAGATG 106179
Qy 1021 CGCTTGGCTGCTGGTGCAACCTTGAAGAAATTTGGTCTGACCCCAAGATAAGATCAAGACC 1080
Db 106180 CGCTTGGCTGCTGGTGCAACCTTGAAGAAATTTGGTCTGACCCCAAGATAAGATCAAGACC 106239
Qy 1081 CACGGTGAGCACTGCACTGCGGATCACACGGAAGATCCAAACAACCGGCTCCGCCCA 1140
Db 106240 CACGGTGAGCACTGCACTGCGGATCACACGGAAGATCCAAACAACCGGCTCCGCCCA 106299
Qy 1141 GATACCGAACTATCACCGGTACCGGTACACGAGCGAGCTGGCGTTTCGCTTGAAGGT 1200
Db 106300 GATACCGAACTATCACCGGTACCGGTACACGAGCGAGCTGGCGTTTCGCTTGAAGGT 106359
Qy 1201 GCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCATGCTGGTGAAATGACC 1260
Db 106360 GCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCATGCTGGTGAAATGACC 106419
Qy 1261 TGGCGTGGTCCGACTTTGAACCTGCTGCTCGTGACACGCGGCTTGGCTGAGTTG 1320
Db 106420 TGGCGTGGTCCGACTTTGAACCTGCTGCTCGTGACACGCGGCTTGGCTGAGTTG 106479
Qy 1321 ACCGTGCTGGTGTGCAACCAATTTGGTTTCTTGGTGCTGCTGCGGGAAGAGGAC 1380
Db 106480 ACCGTGCTGGTGTGCAACCAATTTGGTTTCTTGGTGCTGCTGCGGGAAGAGGAC 106539
Qy 1381 TTCATTTCAAAGCGATGCGCACCGGATTTATCGGCGATATCACACACCTCTCTTCAAGCT 1440
Db 106540 TTCATTTCAAAGCGATGCGCACCGGATTTATCGGCGATATCACACACCTCTCTTCAAGCT 106599
Qy 1441 CCACCTGGGATGATGACAGGGACGATCTGGATTACTTTGGCAGATGTCAACCGTGAAC 1500
Db 106600 CCACCTGCTGATGATGACAGGGACGATCTGGATTACTTTGGCAGATGTCAACCGTGAAC 106659
Qy 1501 AAGCCTCATGTTGCGTCCAAAGAGTGTGCAGACCAATCGATAAGCTGCCCAACATC 1560
Db 106660 AAGCCTCATGTTGCGTCCAAAGAGTGTGCAGACCTCTATCGATAAGCTGCCCAACATC 106719
Qy 1561 AAGATCTGCCACTGCCACCGGTTCCGTTGACCGCTGGAAGAGTGTGGCCCAAGCGCG 1620
Db 106720 AAGATCTGCCACTGCCACCGGTTCCGTTGACCGCTGGAAGAGTGTGGCCCAAGCGCG 106779
Qy 1621 TTTGCTGCTGATCTCCGTGACAGAGCGCACTGGCAGTTACTGATACACACTTCCGCGAT 1680
Db 106780 TTTGCTGCTGATCTCCGTGACAGAGCGCACTGGCAGTTACTGATACACACTTCCGCGAT 106839
Qy 1681 GCACACCACTTTGCTTGGACCCGAGTCCGCTCATTTGCACTGAAGCCTGCGGCAGAG 1740
Db 106840 GCACACCACTTTGCTTGGACCCGAGTCCGCTCATTTGCACTGAAGCCTGCGGCAGAG 106899
Qy 1741 GCCGTGCAAAAGCTGACTCTTGAGCTTTTGTTCGTGAGGCTTGGGGCGGCGCACTTAC 1800
Db 106900 GCCGTGCAAAAGCTGACTCTTGAGCTTTTGTTCGTGAGGCTTGGGGCGGCGCACTTAC 106959
Qy 1801 GATGTGGGATGCGTTTCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGGCGAG 1860
Db 106960 GATGTGGGATGCGTTTCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGGCGAG 107019
Qy 1861 GCGATGCCGAATGAACATTTAGATGCTGCTCGCGGCCCAACACCGTGGGATACACC 1920
Db 107020 GCGATGCCGAATGAACATTTAGATGCTGCTCGCGGCCCAACACCGTGGGATACACC 107079
Qy 1921 CCGTACCAGACTCCGCTGCGCGGCTTTGTTAAGGAAGTTCGACGCTCCGCGGTGGAC 1980
Db 107080 CCGTACCAGACTCCGCTGCGCGGCTTTGTTAAGGAAGTTCGACGCTCCGCGGTGGAC 107139
Qy 1981 ATCTTCCGCACTTTCGAGCGGCTTAAACGACGCTCTCCAGATGCGTCCAGCAATCGACGA 2040
Db 107140 ATCTTCCGCACTTTCGAGCGGCTTAAACGACGCTCTCCAGATGCGTCCAGCAATCGACGA 107199

Qy 2041 GTCTTGAGACCAACACCGCGGTAGCCGAGTGGCTATGGCTTATTCTTGGTGATCTCTCT 2100
Db 107200 GTCTTGAGACCAACACCGCGGTAGCCGAGTGGCTATGGCTTATTCTTGGTGATCTCTCT 107259
Qy 2101 GATCCAAATGAAGAAGCTCTACACCTCGGATTTACTACCTTAAGATGCGAGAGGATCGTC 2160
Db 107260 GATCCAAATGAAGAAGCTCTACACCTCGGATTTACTACCTTAAGATGCGAGAGGATCGTC 107319
Qy 2161 AAGTCTGGCGCTCACATTTCTGGCCATTAAGGATATGGCTTGTCTGCGCCAGCTGCG 2220
Db 107320 AAGTCTGGCGCTCACATTTCTGGCCATTAAGGATATGGCTTGTCTGCGCCAGCTGCG 107379
Qy 2221 GTAAACCAAGTGTGTACCGCACTGCGCGGTGAATTCGATCTGCCAGTGCAAGTGACACC 2280
Db 107380 GTAAACCAAGTGTGTACCGCACTGCGCGGTGAATTCGATCTGCCAGTGCAAGTGACACC 107439
Qy 2281 CAGACACTCGGGTGGCGAGTTGGCTACTTCTGCTGCAGCTCAAGCTGGTGCGAGAT 2340
Db 107440 CAGACACTCGGGTGGCGAGTTGGCTACTTCTGCTGCAGCTCAAGCTGGTGCGAGAT 107499
Qy 2341 GCTGTTGACGCTGCTCCGCAACCACTGCTGGCACCACTCTCCAGCCATCCCTGCTGCGC 2400
Db 107500 GCTGTTGACGCTGCTCCGCAACCACTGCTGGCACCACTCTCCAGCCATCCCTGCTGCGC 107559
Qy 2401 ATTGTTGCTGATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCT 2460
Db 107560 ATTGTTGCTGATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCT 107619
Qy 2461 GACCTCAGCGCTACTGGGAAGCTGTGCGGAGCTGTACTGCTCCATTTGAGTCTGGAACC 2520
Db 107620 GACCTCAGCGCTACTGGGAAGCTGTGCGGAGCTGTACTGCTCCATTTGAGTCTGGAACC 107679
Qy 2521 CCAGGCCCAACCGGTCGCTGTACCGCACGAAATCCAGCGGACAGATTTGCTCAACCTG 2580
Db 107680 CCAGGCCCAACCGGTCGCTGTACCGCACGAAATCCAGCGGACAGATTTGCTCAACCTG 107739
Qy 2581 CGTGCAAGGCCCAACCGCTGCGGCTTTGCTGATCGCTTCGAGCTCATCGAAGACAACCTAC 2640
Db 107740 CGTGCAAGGCCCAACCGCTGCGGCTTTGCTGATCGCTTCGAACTCATCGAAGACAACCTAC 107799
Qy 2641 GCAGCCGTTAATGAGATGCTGGGACCGCCCAACAGAGTCAACCCCATCTCCCAAGGTTGTT 2700
Db 107800 GCAGCCGTTAATGAGATGCTGGGACCGCCCAACAGAGTCAACCCCATCTCCCAAGGTTGTT 107859
Qy 2701 GGCAGCTCGCACTCCACCTGGTTGGTGGCGGTAGATCCAGCAGACTTTGCTGCGAGAC 2760
Db 107860 GGCAGCTCGCACTCCACCTGGTTGGTGGCGGTAGATCCAGCAGACTTTGCTGCGGAT 107919
Qy 2761 CCACAAAGTACGACATCCAGACTCTGTCTCATCGCTTCTGCGCGCGAGCTTTGGTAAC 2820
Db 107920 CCACAAAGTACGACATCCAGACTCTGTCTCATCGCTTCTGCGCGCGAGCTTTGGTAAC 107979
Qy 2821 CTTCCAGGTGGCTGGCCAGAAACCACTCGGCACCCGCGACTGGGAAGCGCTCCGAAGGC 2880
Db 107980 CTTCCAGGTGGCTGGCCAGAGCCACTCGCACCCGCGACTGGGAAGCGCTCCGAAGGC 108039
Qy 2881 AAGSACCTCTGACGGAAGTTCTTGAGAAAGAGCAGCGCACTTCGACGCTGATGATTCC 2940
Db 108040 AAGSACCTCTGACGGAAGTTCTTGAGAAAGAGCAGCGCACTTCGACGCTGATGATTCC 108099
Qy 2941 AAGSACCTCGCAACACGCTCAACCGCTGCTGTTCCGGAAGCCCAACCGAAGAGTTCCTC 3000
Db 108100 AAGSACCTCGCAATAGCTCAACCGCTGCTGTTCCGGAAGCCCAACCGAAGAGTTCCTC 108159
Qy 3001 GAGCACGCTCGCGCTTCGGCAACACCTCTGCGCTGAGATGATCGTGAATTTCTTCAAGGA 3060
Db 108160 GAGCACGCTCGCGCTTCGGCAACACCTCTGCGCTGAGATGATCGTGAATTTCTTCAAGGA 108219
Qy 3061 CTGCTCAGGGCGCGAGACTTTGATCCGCTGCGCATGTCGACCCCACTGCTGTT 3120
Db 108220 CTGCTCAGGGCGCGAGACTTTGATCCGCTGCGCATGTCGACCCCACTGCTGTT 108279


```
Db 3100 GAGCACCGTCGCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATCTTCTACGGC 3159
Qy 3061 CTGGTCGAGGGCGGAGACTTTGATCGCTGCCAGATGCGCAACCCACCTGCTTGT 3120
Db 3160 CTGGTCGAAAGGCGGAGACTTTGATCGCGCTGCCAGATGCGCAACCCACCTGCTTGT 3219
Qy 3121 CGCTCGATGCGATCTCTGAGCAGACGATAAGGCTATGCGCAATCTTGTGGCCAAACGTC 3180
Db 3220 CGCTCGATGCGATCTCTGAGCAGACGATAAGGCTATGCGCAATCTTGTGGCTAACGTC 3279
Qy 3181 AACGGCCAGATCCGCCCAATGCTGCTGACCGCTCCGCTTGAGTCTGTCAACCGCAACC 3240
Db 3280 AACGGCCAGATCCGCCCAATGCTGCTGACCGCTCCGCTTGAGTCTGTCAACCGCAACC 3339
Qy 3241 GCAGAAAGGCGAGATTCCTCAACAAAGGSCCATGCTGCACATTCCTCGTGTGTC 3300
Db 3340 GCAGAAAGGCGAGATTCCTCAACAAAGGSCCATGCTGCACATTCCTCGTGTGTC 3399
Qy 3301 ACTGTGACTGTTGCTGAAGGTGATGAGTCAAGCTCAAGCTGAGATCAGTCCGAATCATCGAG 3360
Db 3400 ACTGTGACTGTTGCTGAAGGTGATGAGTCAAGCTCAAGCTGAGATCAGTCCGAATCATCGAG 3459
Qy 3361 GCTATGAAGATGGAAGCAACATCACTGCTCTGTTGACGCGCAAGATTGAACGCGTTGTG 3420
Db 3460 GCTATGAAGATGGAAGCAACATCACTGCTCTGTTGACGCGCAAAATCGATCGCGTTGTG 3519
Qy 3421 GTTCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCTGCTTCTCTAA 3474
Db 3520 GTTCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCTGCTTCTCTAA 3573

RESULT 10
AX120849
LOCUS AX120849 3420 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 765 from Patent EP1108790.
ACCESSION AX120849
VERSION AX120849.1 GI:14037564
KEYWORDS
SOURCE Corynebacterium glutamicum.
ORGANISM Corynebacterium glutamicum
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Corynebacteriaceae;
Corynebacterium
1. (bases 1 to 3420)
Nakagawa, S., Misoguchi, H., Ando, S., Hayaishi, M., Ochiai, K.,
Yokoi, H., Takeishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
Novel polynucleotides 765 20-JUN-2001;
Patent: EP 1108790-A
KYOMA HAKKO KOGYO CO., LTD. (JP)
Location/Qualifiers
1. 3420
/organism="Corynebacterium glutamicum"
/db_xref="taxon:1718"
BASE COUNT 732 a 1004 c 939 g 745 t
ORIGIN
Query Match 96.3%; Score 3344.8; DB 6; Length 3420;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3373; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 52 GTGTGACTCACACATCTTCAACGCTTCCAGCATTTCAAAAAGATCTTGGTAGCAACCGC 111
Db 1 GTGTGACTCACACATCTTCAACGCTTCCAGCATTTCAAAAAGATCTTGGTAGCAACCGC 60
Qy 112 GCGGAATCGCGTCCGTGCTTCCGTGACGACTCCGAACCGGTGACGACCGTAGCT 171
Db 61 GCGGAATCGCGTCCGTGCTTCCGTGACGACTCCGAACCGGTGACGACCGTAGCT 120
Qy 172 ATTTACCCCGGTGAGATCGGGATCATTCACCGCTCTTTTGTCTGAGCTGCGC 231
Db 121 ATTTACCCCGGTGAGATCGGGATCATTCACCGCTCTTTTGTCTGAGCTGCGC 180
Qy 232 ATTGGTACTGAAGGCTCACCGTCAAGCGGTACCTGGACATCGAATATTACGGTGA 291
```

```
Db 181 ATTGTATCCGGAAGGCTCACAGTCAAGGGCTACCTGGACATCGATGAATATTATCGGTCA 240
Qy 292 GCTAAAAAGTTAAAGCAGATGCTATTATACCCGGGATATGGCTTCTCTGTGAAAAATGCC 351
Db 241 GCTAAAAAGTTAAAGCAGATGCTATTATACCCGGGATATGGCTTCTCTGTGAAAAATGCC 300
Qy 352 CAGCTTGGCCCGCAGTGGCGGAAAAACGCGATTACTTTTATTGGCCCAACCCAGAGGTT 411
Db 301 CAGCTTGGCCCGCAGTGGCGGAAAAACGCGATTACTTTTATTGGCCCAACCCAGAGGTT 360
Qy 412 CTTGATCTCACCGTGATAAAGTCTCGTGGCGTAAACCGCCGCAAGAAGGCTGGTCTGCGCA 471
Db 361 CTTGATCTCACCGTGATAAAGTCTCGCGGGTAAACCGCCGCAAGAAGGCTGGTCTGCGCA 420
Qy 472 GTTTTGGCGGAATCACCCGAGCAAAAACATCGATGACATCGTTTAAAGCGCTGAAGGC 531
Db 421 GTTTTGGCGGAATCCACCCGAGCAAAAACATCGATGACATCGTTTAAAGCGCTGAAGGC 480
Qy 532 CAGCTTACCCCATCTTTTGAAGGCAGTTGCGGGTGGCGGAGCGGATATGCGCTTT 591
Db 481 CAGCTTACCCCATCTTTTGAAGGCAGTTGCGGGTGGCGGAGCGGATATGCGCTTT 540
Qy 592 GTTCTTCACTGATGAGCTCCGCAAAATTGGCAAGAGCATCTCTGMAAGCTGAAGCG 651
Db 541 GTTCTTCACTGATGAGCTTCGCAAAATTAGCAACAAAGACATCTCTGMAAGCTGAAGCG 600
Qy 652 GCATTCGCGCAGCGTTCCGTTATATGTCGAAACGTTGTGATTAAACCCCAAGCATTTGAA 711
Db 601 GCTTTGCGCGATGGCGCGGTATATGTCGAAACGTTGTGATTAAACCCCAAGCATTTGAA 660
Qy 712 GTGCAGATCCTTGGCGGATCGCACTGGAGAAAGTTGTACACCTTTATGAACGTCGCTCA 771
Db 661 GTGCAGATCCTTGGCGGATCGCACTGGAGAAAGTTGTACACCTTTATGAACGTCGCTCA 720
Qy 772 CTGCAAGCTCGTCAACAAAAAGTTGTGAAAAATTGCGCCAGCACAGCATTTGGATCCAGAA 831
Db 721 CTGCAAGCTCGTCAACAAAAAGTTGTGAAAAATTGCGCCAGCACAGCATTTGGATCCAGAA 780
Qy 832 CTGCGTGTATCCCATTTGTGCGGATGCAATAAGTCTGCGCGTCCATTGGTTTACCAGGC 891
Db 781 CTGCGTGTATCCCATTTGTGCGGATGCAATAAGTCTGCGCGTCCATTGGTTTACCAGGC 840
Qy 892 GCGGAACCGTGGAAATCTTGGTCGATGAAAAGGCAACACAGTTTTCATCGAATGAAC 951
Db 841 GCGGAACCGTGGAAATCTTGGTCGATGAAAAGGCAACACAGTCTTTCATCGAATGAAC 900
Qy 952 CCAGTATCCAGTTGAGCACACGCTGACTGAAGAGTCAACGAGGTGAGCTGGGTGAAG 1011
Db 901 CCAGTATCCAGTTGAGCACACGCTGACTGAAGAGTCAACGAGGTGAGCTGGGTGAAG 960
Qy 1012 GCGCAGATGCGCTTGGTGTGCTGGTGCACCTTGAAGAAATTTGGGTCTGACCCAGATGAAG 1071
Db 961 GCGCAGATGCGCTTGGTGTGCTGGTGCACCTTGAAGAAATTTGGGTCTGACCCAGATGAAG 1020
Qy 1072 ATCAAGACCCACGGTGCAGCACCTGCGGATCACCACCGAAGATCCAAACACCGC 1131
Db 1021 ATCAAGACCCACGGTGCAGCACCTGCGGATCACCACCGAAGATCCAAACACCGC 1080
Qy 1132 TTCCGCCCAGATACCGGAATCTATCACCGGTACCGCTCACCGAGGGAGCTGGCGTTCGT 1191
Db 1081 TTCCGCCCAGATACCGGAATCTATCACCGGTACCGCTCACCGAGGGAGCTGGCGTTCGT 1140
Qy 1192 CTTGACGCTGAGCTCAGCTCGGTGGCGAAATCAACGCACTTTTGAATCTCATGCTGCTG 1251
Db 1141 CTTGACGCTGAGCTCAGCTCGGTGGCGAAATCAACGCACTTTTGAATCTCATGCTGCTG 1200
Qy 1252 AAAATGACTCCCGTGTCTCGACTTTGAACATGCTGTTGCTCGTGCAACGCGCGTTG 1311
Db 1201 AAAATGACTCCCGTGTCTCGACTTTGAACATGCTGTTGCTCGTGCAACGCGCGTTG 1260
Qy 1312 GCTGATTCACGCTGCTGCTGTGCAACCAATGGTTTCTGCGTGGCTGGTGGC 1371
```

Db 1261 GCTGAGTTACCGTGTCTGGTGTTCACCAACCAATTGGTTCCTTCCGTGCGTTGCTGCGG 1320
Qy 1372 GAAGAGGACTTCACTTCCOAGCGATGCCACACCGGATTTATCGCGATTCACCCCAACCTC 1431
Db 1321 GAAGAGGACTTCACTTCCOAGCGATGCCACACCGGATTTATTCGCGATTCACCCGACCTC 1380
Qy 1432 CTTGAGGCTCCACCTGCGGATGATGAGGAGGAGCGCATCTCGGATTTACTTGGCAGATGTC 1491
Db 1381 CTTGAGGCTCCACCTGCTGATGATGAGGAGGAGCGCATCTCGGATTTACTTGGCAGATGTC 1440
Qy 1492 ACCGTGAACAAGCCTCATGTGTGCGTCCAAAGGATGTTGCAGACCAACATTCGATAAGCTG 1551
Db 1441 ACCGTGAACAAGCCTCATGTGTGCGTCCAAAGGATGTTGCAGCTCTCTATCGATAAGCTG 1500
Qy 1552 CCCAACATCAAGGATTCGCCACTGCGCAGCGGTTCCCGGTGACCGCTGGAAGCAGCTTGGC 1611
Db 1501 CCTAACATCAAGGATTCGCCACTGCGCAGCGGTTCCCGGTGACCGCTGGAAGCAGCTTGGC 1560
Qy 1612 CCAGCGGTTTGTCTCGTGATCTCGTGAGCAGGACGACCTGGCAGTTACTGTATACCAACC 1671
Db 1561 CCAGCGGTTTGTCTCGTGATCTCGTGAGCAGGACGACCTGGCAGTTACTGTATACCAACC 1620
Qy 1672 TTCCGCGATGCACACCACTCTTTGTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCT 1731
Db 1621 TTCCGCGATGCACACCACTCTTTGTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCT 1680
Qy 1732 GCGCAGAGGCGGTGCGAAAGCTGACTCTCTGAGCTTTTGTCCGTGAGGCGTGGGGCGGC 1791
Db 1681 GCGCAGAGGCGGTGCGAAAGCTGACTCTCTGAGCTTTTGTCCGTGAGGCGTGGGGCGGC 1740
Qy 1792 GCGACCTACGATGTGGGATCGGTTCTCTTTGAGGATCCGTGGGACAGGCTCGACGAG 1851
Db 1741 GCGACCTACGATGTGGGATCGGTTCTCTTTGAGGATCCGTGGGACAGGCTCGACGAG 1800
Qy 1852 CTGCGCAGGCGGATGCGGAATGTAAACATTCAGATGCTGCTTCGCGGCGCAACACCGTG 1911
Db 1801 CTGCGCAGGCGGATGCGGAATGTAAACATTCAGATGCTGCTTCGCGGCGCAACACCGTG 1860
Qy 1912 GGATACACCCCGTACCCAGACTCCGTCGCGCGGCTTTGTTAAGGAAGCTGCCAGCTCC 1971
Db 1861 GGATACACCCCGTACCCAGACTCCGTCGCGCGGCTTTGTTAAGGAAGCTGCCAGCTCC 1920
Qy 1972 GCGGTGACATCTTCCGATCTTCGACGCGTTTAAACGAGTCTCCAGATGCGTCCAGCA 2031
Db 1921 GCGGTGACATCTTCCGATCTTCGACGCGTTTAAACGAGTCTCCAGATGCGTCCAGCA 1980
Qy 2032 ATCGACGAGTCTCGGAGACCAACACGCGGTAGCGAGTGGCTATGGCTTTATCTGCT 2091
Db 1981 ATCGACGAGTCTCGGAGACCAACACGCGGTAGCGAGTGGCTATGGCTTTATCTGCT 2040
Qy 2092 GATCTCTGATCCAAATGAAAGCTCTACACCCCTGGATTTACTACCTAAAGATGGCAGAG 2151
Db 2041 GATCTCTGATCCAAATGAAAGCTCTACACCCCTGGATTTACTACCTAAAGATGGCAGAG 2100
Qy 2152 GAGATGTCAGTCTGGGCTCAATTCGCGCAATTAAGGATATGGCTGCTGCTTCGC 2211
Db 2101 GAGATGTCAGTCTGGGCTCAATTCGCGCAATTAAGGATATGGCTGCTGCTTCGC 2160
Qy 2212 CCAGCTCGGTAAACCAAGCTGCTCACGCGCTGCGCGGTGAATTCGATCTGCCAGTGCAC 2271
Db 2161 CCAGCTCGGTAAACCAAGCTGCTCACGCGCTGCGCGGTGAATTCGATCTGCCAGTGCAC 2220
Qy 2272 GTGCACCCACGACACTGCGGGTGGCGAGTTGGCTACTACTTTGTGTCGAGCTCAAGCT 2331
Db 2221 GTGCACCCACGACACTGCGGGTGGCGAGTTGGCTACTACTTTGTGTCGAGCTCAAGCT 2280
Qy 2332 GGTGCGATGCTGTTGAGGCTGCTTCGCGACCACTGCTGCGCAACACTCCCGACCAATCC 2391
Db 2281 GGTGCGATGCTGTTGAGGCTGCTTCGCGACCACTGCTGCGCAACACTCCCGACCAATCC 2340
Qy 2392 CTGCTGCCATTTGTTGCTGCTTCGCGACACACCGTTCGCGATACCGGTTTGAGCCTCGAG 2451
Db 2341 CTGCTGCCATTTGTTGCTGCTTCGCGACACACCGTTCGCGATACCGGTTTGAGCCTCGAG 2400

RESULT 11
AB083299
LOCUS

AB083299

4851 bp

DNA

linear

BCT 10-APR-2002

Qy 2452 GCTGTTTCTGACCTCGAGCCGTACTTGGGAAAGCTGTGCGGACTGTACCTGCCATTTGAG 2511
Db 2401 GCTGTTTCTGACCTCGAGCCGTACTTGGGAAAGCTGTGCGGACTGTACCTGCCATTTGAG 2460
Qy 2512 TCTGGAAACCCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCAGGCGGACAGTTG 2571
Db 2461 TCTGGAAACCCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCAGGCGGACAGTTG 2520
Qy 2572 TCCAACCTGGGTGCACAGGCCACCGCACTGGGCTTCTGATCTGCTTCGAGCTCATCGAA 2631
Db 2521 TCCAACCTGGGTGCACAGGCCACCGCACTGGGCTTCTGCGATCGTTTCGAACCTCATCGAA 2580
Qy 2632 GACAACTAGAGCCGTTAATGAGATGCTGGGAGCGCCAAACCAAGGTCAACCCCATCTCTCC 2691
Db 2581 GACAACTAGAGCCGTTAATGAGATGCTGGGAGCGCCAAACCAAGGTCAACCCCATCTCTCC 2640
Qy 2692 AAGGTTGCTTGGCGACCTCGCACTCCACTGCTGTTGGTGGGCTGTAGATCCAGAGACTTT 2751
Db 2641 AAGGTTGCTTGGCGACCTCGCACTCCACTGCTGTTGGTGGGCTGTAGATCCAGAGACTTT 2700
Qy 2752 GCTGCAGACCCACAAAAGTACGACATCCAGACTCTCTGATCGGCTTCCTGCGCGGCGAG 2811
Db 2701 GCTGCAGATCCACAAAAGTACGACATCCAGACTCTCTGATCGGCTTCCTGCGCGGCGAG 2760
Qy 2812 CTTGGTTAACCTCCAGGTGCTGGCCAGAAACCACTGCGCAACCCCGCACTGGAAGGCGC 2871
Db 2761 CTTGGTTAACCTCCAGGTGCTGGCCAGAAACCACTGCGCAACCCCGCACTGGAAGGCGC 2820
Qy 2872 TCCGAGGCAAGCAGCTCTGAGGAGTTCTTCCAGAAAGCAGGCGCACTCTCGAGCT 2931
Db 2821 TCCGAGGCAAGCAGCTCTGAGGAGTTCTTCCAGAAAGCAGGCGCACTCTCGAGCT 2880
Qy 2932 GATGATTCGAAGAACTCGCAACAGCTCAACCGCTGCTGTTCCCGAAAGCCAAACCGAA 2991
Db 2881 GATGATTCGAAGAACTCGCAACAGCTCAACCGCTGCTGTTCCCGAAAGCCAAACCGAA 2940
Qy 2992 GAGTTCTTCGAGCACCGTCCCGCTTCGCGCAACCACTCTCGGCTGGATGATCGTGAATTC 3051
Db 2941 GAGTTCTTCGAGCACCGTCCCGCTTCGCGCAACCACTCTCGGCTGGATGATCGTGAATTC 3000
Qy 3052 TTCTACGCACTGTTCGAGGCGCGGAGACTTTGATCCGCTGCCAGATGTCGCGCACCCCA 3111
Db 3001 TTCTACGCACTGTTCGAGGCGCGGAGACTTTGATCCGCTGCCAGATGTCGCGCACCCCA 3060
Qy 3112 CTGCTTCTCCCTGGATGCGATCTCTGAGCAGACATTAAGGATATGCGCAATGTTGTG 3171
Db 3061 CTGCTTCTCCCTGGATGCGATCTCTGAGCAGACATTAAGGATATGCGCAATGTTGTG 3120
Qy 3172 GCCAACGTCACGCGCCAGATCCCGCCCAATCGTGTGCGTGAACCGCTCCGTTGAGTCTGTC 3231
Db 3121 GCCAACGTCACGCGCCAGATCCCGCCCAATCGTGTGCGTGAACCGCTCCGTTGAGTCTGTC 3180
Qy 3232 ACCGCAACCGCAGAAAGGAGGATTCCTCCAAACAGGGCCATGTTGCTGCACCACTTCGCT 3291
Db 3181 ACCGCAACCGCAGAAAGGAGGATTCCTCCAAACAGGGCCATGTTGCTGCACCACTTCGCT 3240
Qy 3292 GGTGTTGTCATGCTGATGCTGTTGAGGATGATGAGGTCAAGGCTGAGATGCGAGTCGCA 3351
Db 3241 GGTGTTGTCATGCTGATGCTGTTGAGGATGATGAGGTCAAGGCTGAGATGCGAGTCGCA 3300
Qy 3352 ATCATCGAGGCTATGAGATGGAAGCAACAAATCACTGCTTCTGTTGACGGCAAGATTGAA 3411
Db 3301 ATCATCGAGGCTATGAGATGGAAGCAACAAATCACTGCTTCTGTTGACGGCAAAATCGAT 3360
Qy 3412 CGGTTGTGTTCTGCTGCAACGAGGTGGAAGGTGGCGACTTGATCGTCTGCTGTTCC 3471
Db 3361 CGGTTGTGTTCTGCTGCAACGAGGTGGAAGGTGGCGACTTGATCGTCTGCTGTTCC 3420

DEFINITION

Corynebacterium efficiens pyc gene for pyruvate carboxylase, complete cds.

ACCESSION

AB083299

VERSION

AB083299.1

KEYWORDS

Corynebacterium efficiens DNA.

SOURCE

Corynebacterium efficiens

ORGANISM

Corynebacterium

REFERENCE

1. Akiyoshi, N., Nonaka, G., Kimura, E., Kawahara, Y. and Sugimoto, S.

AUTHORS

Corynebacterium efficiens

TITLE

Corynebacterium efficiens pyruvate carboxylase (pyc) gene, complete

JOURNAL

CDS

REFERENCES

2 (bases 1 to 4851)

TITLE

Published Only in Database (2002)

JOURNAL

Nonaka, G., Akiyoshi, N., Kimura, E., Kawahara, Y. and Sugimoto, S.

AUTHORS

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

TITLE

Direct Submission

JOURNAL

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

AUTHORS

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

TITLE

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

JOURNAL

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

AUTHORS

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

TITLE

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

JOURNAL

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

AUTHORS

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

TITLE

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

JOURNAL

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

AUTHORS

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

TITLE

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

JOURNAL

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

AUTHORS

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

TITLE

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

JOURNAL

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

AUTHORS

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

TITLE

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

JOURNAL

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

AUTHORS

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

TITLE

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

JOURNAL

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

AUTHORS

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

TITLE

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

JOURNAL

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

AUTHORS

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

TITLE

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

JOURNAL

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

AUTHORS

Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,

FEATURES

source

gene

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

FEATURES

source

gene

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

FEATURES

source

gene

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

QY 1259 CCTGCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGACAGCGCGGTTGGCTGAGT 1318
Db 2364 CCTGCGCGGTTCCGATTTTCGAGACCGCGGTGTCCTCGAGCGAGCGCCCTGCGCGAGT 2423
QY 1319 TCACCGTGTCTGGTGTTCACAACTGATGTTTCTTCGTCGCTGTTGCTGCGGGAAGAGG 1378
Db 2424 TCAACGTCTCCGCGTGGCCACCAACATCGCTTCTCTGCTGCTGCTGCGCGAGGAG 2483
QY 1379 ACTTCACTTCCAAAGGCATCGCCACCGGATTTATCGGCGATCAACCCACACTCTCTTCAAG 1438
Db 2484 ACTTCAACAAAGGCGCATCGACACCGGCTTCATCGGCTCCACACAGCACCTGCTCCAGG 2543
QY 1439 CTCCACCTCGCGATGATGAGCAGGAGCATCTCTGGAATTAATTGGCAGATGTCAACGTGA 1498
Db 2544 CCCACCGCGCGACGATGAGCAGGCGCGGATCTCTGGAATACCTGGCGGATGTCAACGTGA 2603
QY 1499 ACAAGCCTCATGCTGTGCTGCTCAAAAGGATGTTGACGACCAATCGATAAGCTGCCAACA 1558
Db 2604 ACAACCCACCGTGAACGCC---CGAGACAGCCCGTCCGATAGAGGAGCTGCCGAGG 2660
QY 1559 TCAAGGATCTGCCACTGCGCA CGCGGTTCCCGTGACCGGCTGAAGCAGCTTGGCCGAGCGG 1618
Db 2661 TGGAGAACATCCCGCTGCGACCGCGCTCCCGGACCGCTGAAGCAGCTCTGGCCGAGG 2720
QY 1619 CGTTTGTCTGATCTCCGTGAGCAGGACGACTGGCAGTTACTGATACACCTTCGCG 1678
Db 2721 GTTTTGGCCCGGATCTGCGCGAAACAGGATGCCCTGGCGCTCACCGACACCACTTCGCG 2780
QY 1679 ATGACACCAAGCTTTGCTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTCGCGGAG 1738
Db 2781 ATGCCACCACTCCTCTGCGCCACCGCGTGGCTGCGCTTCTTCGCGTGACCCCGCGGCGC 2840
QY 1739 AGGCGTGCAGAAAGTGAATCCTGAGCTTTTGTCCGTGAGCGCTTGGGCGCGCGACCT 1798
Db 2841 CGCGCGTCGCAAAAGTCAACCCCGAGCTGCTGCTGGTGAAGCCCTGGGCGGTCGCACCT 2900
QY 1799 ACGATGTGGCGATGGTTTCTCTTTGAGGATCCGTGGGACAGGCTCGACAGCTGGCG 1858
Db 2901 ACGACGTGGCGATGCGCTTCTCTTCGAGGATCCGTGGGACAGGCTGATGAGCTGGTG 2960
QY 1859 AGGCGATGCCAAATGTAAACATTCAAGTGTGCTTTCGGCGCGCAACACCGTGGGATACA 1918
Db 2961 AGGCGATGCCAAATGTGNAATCCAGATGCTGCTGCGTGGTTCGCAACACCGTCCGGTACA 3020
QY 1919 CCCGATACCCAGACTCCGTCTGCGCGCGTTTGTAAAGGAAGCTGCCAGCTCCGCGGTGG 1978
Db 3021 CCCGATACCCGATTCGCTGTGCGCGGCTTGTGTCAGGAGGCGCCAAAGTCCGCTGTGG 3080
QY 1979 ACATCTTCCGATCTTCGACCGGCTTAACGAGCTCTCCAGATGGCTCCAGCAATCGACG 2038
Db 3081 ACATCTTCCGATCTTCGACCGCTCAACGACATCTCCAGATGGCGCGGCCATCGACG 3140
QY 2039 CAGTCTTGGAGACAAACACCGGCTAGCGAGGTGGCTATGGCTTATTTCTGCTGATCTCT 2098
Db 3141 CCGTCTTGGAGACCGGACCAAGTGTGCGAGGTCGCCATGGCTACTTCCGTTGACCTGT 3200
QY 2099 CTGATCCAAATGAAAAGCTCTACACCCCTGGATTACTACTTAAAGATGGCAGAGGATCG 2158
Db 3201 CCAATCCGGGGAGAGCTCTACACCCCTGGACTACTACTTGAACCTTGGCGAGCAGATCG 3260
QY 2159 TCAAGTCTGGCGCTCACATTTCTGGCCATTAAGGATATGGCTGTGCTGTTGCGCCAGCTG 2218
Db 3261 TCGACTCCGGTGCACATCTCTGGCCATCAAGGACATGGCGCGGCTGCTGCGCGCGCGG 3320
QY 2219 CGGTAAACCAAGCTGCTCACCGCACTGCGCGTGAATTCGATCTGCCAGTGCACGTGACACA 2278
Db 3321 CGGCGCCAAACTGCTCACCGCCCTGCGCGTGAATTCGACTGCCCGTGCATGTCCACA 3380
QY 2279 CCCACGACACTGCGGGTGGCCAGTTGGCTACTTACTTTGCTGACAGCTCAAGCTGGTGCAG 2338
Db 3381 CCCACGACACCGCGCGGTGAGCTGGCCACTTACTTGGCGCGCGCCNAACCGCGGCGCG 3440
QY 2339 ATGCTGTGACGGTGTTCGCAACAATGCTGTGGCAACAACCTCCACGCCATCCCTGTCTG 2398

Db 3441 ATGCGCTGACGCGCGCTCCGACCCCTGTCCGGTACACCTCCAGCGCTCGATGTCG 3500
QY 2399 CCATTTGTTGCTGCAATTCGCGCACACCCGTCGCGATACCGGTTTTCGAGCTTCGAGGCTGTTT 2458
Db 3501 CTCTGGTTGCGCGGTTTTCGCGCACACCCGACGCGACACCGGCTCAACCTGACGCGCGTCT 3560
QY 2459 CTGACCTCGAGCCGCTGCTGGGAAGCTGTGCGCGAGCTGTACCTGCCATTTGAGTCTGGAA 2518
Db 3561 CCGACCTGGAAACCGTACTGGGAGGCGGTTCGCGGACTGTACTCTCCGTTTGAATCCGCA 3620
QY 2519 CCCAGGCCCCAACCGGTCGCTTACCGCACCAAGAAATCCAGGCGGACAGTTGTCCAAACC 2578
Db 3621 CCCGCGGCCGACCGGACCGGCTTACCGCCACAGATCCCGCGCGGTGAGTGTCCAAACC 3680
QY 2579 TCGTGTACACAGGCCACCGCACTGGGCTTGTGATCGCTTCGAGCTCATCGAAGACAAC 2638
Db 3681 TCGTGTCCAGGCGGTTGCACTGGGTCTGGCGGACCGTTCGAGCTCATCGAGGACTACT 3740
QY 2639 ACGCAGCCGTTAATGAGATGCTGGGACCGCCAAACCAAGGTCAACCCATTCCTCCAAAGTTG 2698
Db 3741 ACGCGGCGTCAACGAGATGCTGGGTCTCGACCAAGGTCAACCCGCTCTCCAAAGTTG 3800
QY 2699 TTGCGGACCTCGCACTCCACTGGTTGGTGGGGGTGTAGATCCAGCAGACTTTTCTCTCAG 2758
Db 3801 TCGGTGACCTCGCACTGCACCTCGTGGTGGCGGTGTGAGCCCGGAGGATTTCCGCCGCG 3860
QY 2759 ACCCAAAAAGTACGACATCCAGACTCTGTCTATCGCGTTCTTCGCGCGGAGCTTGTA 2818
Db 3861 ATCGCAGAGTACGACATCCCGATTCGCTTCGCTTCCTCCGCGGAGACTGGTA 3920
QY 2819 ACCTTCAGTGGGTGCGCAGAAACCACTGCGCACCCCGGCACTGGAAGGCGCTCCGAAG 2878
Db 3921 CCCCTCCCGTGGCTGCGCCGAAACCGCTGCGCACCCGTCGCACTCGAGGCTCGCTCCGAG 3980
QY 2879 GCAGGCACTCTGACGGAAGTTCTTGAGGAGAGCAGGCGCACCTCGACGCTGATGATT 2938
Db 3981 GTAAGGCCCCGCTGGCGGAGATCCCGCCGAGGAGAGGCCACCTGGATTTCGATGATT 4040
QY 2939 CCAAGAACGTCGCAACAGCCTCAACCGCCTGCTGTTCCCGAAGCCAAACCGAAGTTCC 2998
Db 4041 CCGCGAGAGCTCGCGGACCCCTCAACCGCTGCTGTTCCCGAAGCCGACCGAGGTTCC 4100
QY 2999 TCGAGCAACGTCGCGCTCGGCAACACCTCTGCGCTGGATGATCGTGAATTCCTTCAAG 3058
Db 4101 TTGAGCACCGTCGCGCTTCGCGCAACACCTCGCCCTGGATGACCGCGAGTTCTTCTAAG 4160
QY 3059 GACTGGTTCGAGGCGCGGAGACTTTGATCCGCTTCGAGATGTGCGCACCCCACTGCTTG 3118
Db 4161 GCTTGAAGGAGGAGCGTGGAGAGCTGATCCGACTGACCGGTGTGTCACCCCGATGGTG 4220
QY 3119 TTGCGCTTGGATGCGATCTCTGAGCCAGACGATAAGGATATGCGCAATGTTGTGGCCAAAG 3178
Db 4221 TCAGCTTGGATGCGGTGTCGAAACCGGATGACAAAGGATGCGGCAACGTTGGTGTCAAG 4280
QY 3179 TCAACGCGCAGATCCGCCCAATGCTGTGCGTGAACCGCTCCGTTGAGTCTGTCAACGCA 3238
Db 4281 TCAACGCGCAGATCCGCCCGCTCAAGGTGCGGACCGTTCCGTTGAGTTCGCTACCCGCA 4340
QY 3239 CCGCAGAAAAGGCGAGATTCTCCAAAGGCGCATTTGCTGCAACCAATTCGCTGGTGTG 3298
Db 4341 CCGCGGAGAAAGGCGGATGCCAACCAAGGCGCATGTGCGCGCACCAATTCGCGGTGTG 4400
QY 3299 TCACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGAGATGCAAGTGCATCATCG 3358
Db 4401 TCACCTGACCGTCCGCGGAGGTGATGAGATCAAGGCTGGCGACCGCTGGCCATCATG 4460
QY 3359 AGGCTATGAAGTGAAGCAACAATCACTGCTTCTGTTGACGGCAAGATTGAACGCGTTG 3418
Db 4461 AGGCTATGAAGTGAAGCCCAACCATCAACCGGCTGTGCGACGGTGTGTCATCGACCGCTG 4520
QY 3419 TGGTTCTGCTGCAACGAAAGGTGAAGGTGGGACTTGTGATCGTGTGTTTCTTA 3473

Db 4521 TGGTCCCGCCGCCACCAAGGTCGAGGCGCGGACCTCATCGTGGTGGTGCCTA 4575

RESULT 12
AX064993
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS

AX064993
Sequence 119 from Patent WO0100844.
AX064993
AX064993.1 GI:12542705
Corynebacterium glutamicum.
Corynebacterium glutamicum.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Corynebacteriaceae;
Corynebacterium
1 (bases 1 to 1719)
Pompejus, M., Kroeger, B., Schroeder, H., Zelder, O. and Haberhauer, G.
Corynebacterium glutamicum genes encoding proteins involved in
carbon metabolism and energy production
Patent: WO 0100844-A 119 04-JAN-2001;
BASF AKTIENGESELLSCHAFT (DE)
Location/Qualifiers
1..1719
/organism="Corynebacterium glutamicum"
/db_xref="taxon:1718"
101..1699
/note="RXN02328"
/codon_start=1
/transl_table=1
/protein_id="CAC25736.1"
/db_xref="GI:12542706"
/translation="MTAITLGLLLKGIITVSTHTSLPAPFKKILVANRGEIAVRA
FRAALTEGAATVAIYPREDGSRPHRSFASAEVRIETGSPVKAYLIDIEITGAARKVK
ADATYPGYFLSENAQLARECAENGITFIPTPEVLDTGDKSRVAVTAAKKAGIPLVIA
ESTSKNIDETIVKSAEQTYPIFKVAVAGGGRGMRFVASDELKFLATEASREAEA
PDGAVYVERAVINPOHIEVOIGLDHDTGEVHLRYERDSCLORRHOKVVEIAPADHLP
ELRDICADAVKFCRSIGYOGAGTVEFLDEKGNHVFIMNPRIOEHVTEVEYED
LVKQMLRCAAGATLKEGLQDKIKTHGAALOCRIITTEDPNRPRDPTGTTAVRSPG
GAGVLDLCAALQGEITAHFDSMLVMKTCRQSDFFETAVARAORALAEFTVSVATNIG
FLRALREEDFTSKRIATGFIADHPHLLQAPADDEQGRILDYLDVADTVNPKHVRPK
DVAAPDKLPNIKDLPLPRGSRDR"
BASE COUNT 405 a 457 c 468 g 389 t
ORIGIN

Query Match 45.0%; Score 1564.8; DB 6; Length 1719;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 1592; Conservative 0; Mismatches 27; Indels 1; Gaps 1;

Qy 1 GTGACTGCTATCACCTTGGCGGTCTCTGTGGAAGGATAATTACTTAGTGTGACT 60
Db 101 GTGACTGCTATCACCTTGGCGGTCTCTGTGGAAGGATAATTACTTAGTGTGACT 160
Qy 61 CACATCTTCAACGGTTCAGCATTCAAAAGATCTTGGTAGCAACCGCGCGGAAATC 120
Db 161 CACATCTTCAACGGTTCAGCATTCAAAAGATCTTGGTAGCAACCGCGCGGAAATC 220
Qy 121 GCGGTCCGTGCTTTCCGTGAGCATTCGAAACCGGTGAGCCAGCGTAGCTATTTACCCC 180
Db 221 GCGGTCCGTGCTTTCCGTGAGCATTCGAAACCGGTGAGCCAGCGTAGCTATTTACCCC 280
Qy 181 CGTGAAGATCGGGGATCAATTCACCGCTCTTTTGTCTGAAGCTGTCGCCATTTGACT 240
Db 281 CGTGAAGATCGGGGATCAATTCACCGCTCTTTTGTCTGAAGCTGTCGCCATTTGACT 340
Qy 241 GAAGGCTCACAGTCAAGCGGTACCTGGACATCGATGAATTTATCGGTGAGCTAAATA 300
Db 341 GAAGGCTCACAGTCAAGCGGTACCTGGACATCGATGAATTTATCGGTGAGCTAAATA 400
Qy 301 GTTAAAGCAGATGCTATTTACCGGGATATGCTTCTGTCTGAAATGCCAGCTTGCC 360
Db 401 GTTAAAGCAGATGCTATTTACCGGGATATGCTTCTGTCTGAAATGCCAGCTTGCC 460
Qy 361 CGCGAGTGGCGGAAACGGCATTAATTTATGGCCCAACCCGAGGTTCTTGATCTC 420

Db 461 CCAGAGTGTGCGGAAAACGGCATTTACTTTATTGGCCCAACCCAGAGGTTCTTGATCTC 520
Qy 421 ACCGCTGATAAGTCTCGTGGCGGTAAACCGCGCGGAGAAAGGCTGGTCTGCCAGTTTGGCG 480
Db 521 ACCGCTGATAAGTCTCGCGGGTAAACCGCGCGGAGAAAGGCTGGTCTGCCAGTTTGGCG 580
Qy 481 GAATCCACCCCGAGCAAAAAACATCGATGACATCGTTTAAAGCGCTGAAGGCCAGACTTAC 540
Db 581 GAATCCACCCCGAGCAAAAAACATCGATGAGATCGTTTAAAGCGCTGAAGGCCAGACTTAC 640
Qy 541 CCATCTTTTGTAAAGGAGTGTGCGGTGGTGGCGGACGCGGTATGCGTTTGTCTTCA 600
Db 641 CCATCTTTTGTAAAGGAGTGTGCGGTGGTGGCGGACGCGGTATGCGTTTGTCTTCA 700
Qy 601 CCTGATGATCCCGCAAAATTTGGCAACAGAGCATCTCGTGAAGCTGAAGCGCATTTCCGC 660
Db 701 CCTGATGATCCCGCAAAATTTAGCAACAGAGCATCTCGTGAAGCTGAAGCGCATTTCCGC 760
Qy 661 GACGTTTCGGTATATGTGCAACGCTGTGATTAACCCCGACACATCGAAGTGCAGATC 720
Db 761 GATGCGCGGTATATGTGCAACGCTGTGATTAACCCCGACATATCGAAGTGCAGATC 820
Qy 721 CTTGGCGATCCGATGAGAGAGTTGTACACCTTTATGAACGCTGCTCACTGCGAGGT 780
Db 821 CTTGGCGATCACCTGAGAGAGTTGTACACCTTTATGAACGCTGCTCACTGCGAGGT 880
Qy 781 CQTCAACAAAAGTTGTGCAAAATTTGCGCCAGCACAGCATTTGGATCCAGAACTGCGGTAT 840
Db 881 CQTCAACAAAAGTTGTGCAAAATTTGCGCCAGCACAGCATTTGGATCCAGAACTGCGGTAT 940
Qy 841 CGCATTTGTGCGGATGCAAGAAATTTCTGCGCTCCATTTGGTTTACAGGCGCGGGAACC 900
Db 941 CGCATTTGTGCGGATGCAAGAAATTTCTGCGCTCCATTTGGTTTACAGGCGCGGGAACC 1000
Qy 901 GTGGAATTTCTTGTGCTGATGAAAAGGGCAACACGTTTTCATCGAAATGAACCCAGCTATC 960
Db 1001 GTGGAATTTCTTGTGCTGATGAAAAGGGCAACACGCTTTCATCGAAATGAACCCAGCTATC 1060
Qy 961 CAGGTTTGAAGCACACCGTGACTGAAAGAGTCAACCGAGTGGACCTGCTGAAAGCGCAGATG 1020
Db 1061 CAGGTTTGAAGCACACCGTGACTGAAAGAGTCAACCGAGTGGACCTGCTGAAAGCGCAGATG 1120
Qy 1021 CGCTTGGCTGTGGTGGCAACCTTTGAAGGAATTTGGGTCTCACCAAGATAAGATCAAGACC 1080
Db 1121 CGCTTGGCTGTGGTGGCAACCTTTGAAGGAATTTGGGTCTCACCAAGATAAGATCAAGACC 1180
Qy 1081 CACGTTGCGACACTGCACTGCCGTCACCGAGATCCACCAAGAGATCCAAAACAAAGGCTTCCGCCA 1140
Db 1181 CACGTTGCGACACTGCACTGCCGTCACCGAGATCCACCAAGAGATCCAAAACAAAGGCTTCCGCCA 1240
Qy 1141 GATACCGGAATATATCACCGCTAGCGCTCACAGGCGGAGCTGGGCTTCTTGAAGGT 1200
Db 1241 GATACCGGAATATATCACCGCTAGCGCTCACAGGCGGAGCTGGGCTTCTTGAAGGT 1300
Qy 1201 GAGCTCAGCTCGGTGGCGAATAACACCGCAACATTTGATCTCATGCTGGTGAATAATGACC 1260
Db 1301 GAGCTCAGCTCGGTGGCGAATAACACCGCAACATTTGATCTCATGCTGGTGAATAATGACC 1360
Qy 1261 TGCCTGGTTCGCACTTTGAAACTGTCTTGTCTGCTGCAAGCGCGCTTGGCTGAGTTTC 1320
Db 1361 TGCCTGGTTCGCACTTTGAAACTGTCTTGTCTGCTGCAAGCGCGCTTGGCTGAGTTTC 1420
Qy 1321 ACCGCTGCTGGTGTGCAACCAACATTTGGTTTCTTGTGCTGCTGCTGCGGGAAGAGGAC 1380
Db 1421 ACCGCTGCTGGTGTGCAACCAACATTTGGTTTCTTGTGCTGCTGCTGCGGGAAGAGGAC 1480
Qy 1381 TTCACTTCCAGCGCATCGCCACCGGATTTATCGCGGATCACCCACACCTCTCTCAGGCT 1440
Db 1481 TTCACTTCCAGCGCATCGCCACCGGATTTATCGCGGATCACCCACACCTCTCTCAGGCT 1540
Qy 1441 CCACCTCGCGGATGATGAGCGGAGCGATCCTGGATTACTTTGGCAGATCTGACCGTGAAC 1500

Db 1541 CCACCTGCTGATGATGACGAGGAGCCATCTCGATTACTTGGCAGATGTCACCGTGAAC 1600

Qy 1501 AAGCCTCATGTGTCGTCGTCACAAAGATGTCAGACCAATCGATAAGCTGCCCAACATC 1560

Db 1601 AAGCCTCATGTGTCGTCGTCACAAAGATGTCAGATGTCATCGATTAAGCTGCCTAACATC 1660

Qy 1561 AAGGATCTGCCACATGCCACGCGGTTCCCGTCACCGCTGAAGCAGCTTGGCCGACCGCG 1620

Db 1661 AAGGATCTGCCACATGCCACGCGGTTCCCGTGACCG-CTGAAGCAGCTTGGCCGACCGCG 1719

RESULT 13

AX064995

LOCUS AX064995 1406 bp DNA linear PAT 24-JAN-2001

DEFINITION Sequence 121 from Patent WO0100844.

ACCESSION AX064995

VERSION AX064995.1 GI:12542707

KEYWORDS

SOURCE Corynebacterium glutamicum.

ORGANISM Corynebacterium glutamicum

Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE 1 (bases 1 to 1406)

AUTHORS Pompejus,M., Kroeger,B., Schroeder,H., Zelder,O. and Haberhauer,G.

TITLE corynebacterium glutamicum genes encoding proteins involved in carbon metabolism and energy production

JOURNAL Patent: WO 0100844-A 121 04-JAN-2001;

BASF AKTIENGESSELLSCHAFT (DE)

FEATURES

source 1..1406

/organism="Corynebacterium glutamicum"

/db_xref="taxon:1718"

<1..1386

/note="FRXA02328"

/codon_start=1

/transl_table=11

/protein_id="CAC25737.1"

/db_xref="GI:12542708"

/CDS

BASE COUNT 325 a 381 c 385 g 315 t

ORIGIN

Query Match 38.9%; Score 1351.8; DB 6; Length 1406;

Best Local Similarity 98.0%; Pred. No. 1.3e-280;

Matches 1379; Conservative 0; Mismatches 27; Indels 1; Gaps 1;

Qy 214 GCTTCTGAAGCTGCCGATGTTGTTGATGCTCACCAGGCTCACCAGTCAAGGCGTACCTGGACATC 273

Db 1 GCCTTCTGAAGCTGCCGATGTTGTTGATGCTCACCAGGCTCACCAGTCAAGGCGTACCTGGACATC 60

Qy 274 GATCAATATTATCGGTGAGCTTAAAGTTAAAGCAGATGCTATTACCGGGATATGGC 333

Db 61 GATCAATATTATCGGTGAGCTTAAAGTTAAAGCAGATGCTATTACCGGGATATGGC 120

Qy 334 TTCCTGTCTGAAATGCCAGCTTCCCGCGAGTGCGGAAACCGCATTTTATT 393

Db 121 TTCCTGTCTGAAATGCCAGCTTCCCGCGAGTGCGGAAACCGCATTTTATT 180

Qy 394 GGCCCAACCCAGAGGTTCTTGATCTCACCAGGCTGATAGTCTCGTGGGGTAACCGCGCG 453

Db 181 GGCCCAACCCAGAGGTTCTTGATCTCACCAGGCTGATAGTCTCGCGGGTAACCGCGCG 240

Qy 454 AAGAGGCTGCTGCCAGTTTGGCGAATCCACCCGAGCAAAACATCGATGAATC 513

Db 241 AAGAGGCTGGTCTGCCAGCTTTTGGCGGAATCCACCCGAGCAAAAAATCATCGATGAGATC 300

Qy 514 GTTAAAAAGCGCTGAAGCCAGACTTACCCCATCTTTGTAAAGGAGTTCGCGGTGGTGGC 573

Db 301 GTTAAAAAGCGCTGAAGCCAGACTTACCCCATCTTTGTGAAGGAGTTCGCGGTGGTGGC 360

Qy 574 GGAGCGGATATGCGCTTTGTTTCTTACCTGATGAGCTCCGCAAAATGGCAACAGAGCA 633

Db 361 GGAGCGGATATGCGCTTTGTTTCTTACCTGATGAGCTTCGCAAAATAGCAACAGAGCA 420

Qy 634 TCTCGTGAAGCTGAAGCGCATTTCCGCGACGGTTCGGGTATATATGTCGAACGCTGCTGATT 693

Db 421 TCTCGTGAAGCTGAAGCGCATTTCCGCGATGCGCGGTATATGTCGAACGCTGCTGATT 480

Qy 694 AACCCCCAGACATGTAAGTGCAGATCTTTGGCGATCGCATCGGAGAGTGTGACACCTT 753

Db 481 AACCTCAGCATATTGAAGTGCAGATCTTTGGCGATCACACTGGAGAAGTTGTACACCTT 540

Qy 754 TATGAAGCTGACTGCTCACTGCAGCTGTCACCAAAAAAGTTGTCGAAATTCGCCACGA 813

Db 541 TATGAAGCTGACTGCTCACTGCAGCTGTCACCAAAAAAGTTGTCGAAATTCGCCACGA 600

Qy 814 CAGCATTTGATCCAGAACTGCGTGCATTTGTGCGGATGCAGTAAAGTTCTGCGCG 873

Db 601 CAGCATTTGATCCAGAACTGCGTGCATTTGTGCGGATGCAGTAAAGTTCTGCGCG 660

Qy 874 TCCATTGTTTACAGGCGCGGAAACCGTGGAAATTCCTGTCGATGAAAAAGGCAACAC 933

Db 661 TCCATTGTTTACAGGCGCGGAAACCGTGGAAATTCCTGTCGATGAAAAAGGCAACAC 720

Qy 934 GTTTTCATCGAAATGAACCAACGATATCCAGGTTGAGCACACCGTGCATGAAAGATGACC 993

Db 721 GTCTTCATCGAAATGAACCAACGATATCCAGGTTGAGCACACCGTGCATGAAAGATGACC 780

Qy 994 GAGTGGACCTGTTGAGGCGCAGATGCGCTGCTGCTGTCGTCGCAACCTTGAAGGAATTG 1053

Db 781 GAGTGGACCTGTTGAGGCGCAGATGCGCTGCTGCTGTCGTCGCAACCTTGAAGGAATTG 840

Qy 1054 GGTCTGACCCCAAGATGAAGATCAAGACCCACGGTGCAGCACTGCGATGCCGATCACCACG 1113

Db 841 GGTCTGACCCCAAGATGAAGATCAAGACCCACGGTGCAGCACTGCGATGCCGATCACCACG 900

Qy 1114 GAGATCAAAACAAACGGCTTCCGCCACGATACCGGAACATATACCGCGTACCGCTACCA 1173

Db 901 GAGATCAAAACAAACGGCTTCCGCCACGATACCGGAACATATACCGCGTACCGCTACCA 960

Qy 1174 GCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGTGGCGGAAATCACCAGCAC 1233

Db 961 GCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGTGGCGGAAATCACCAGCAC 1020

Qy 1234 TTTGACTCCCATGCTGTTGAAAAATGACCTGCGGTGGTTCCGACTTTGAAACTGCTGTTGCT 1293

Db 1021 TTTGACTCCCATGCTGTTGAAAAATGACCTGCGGTGGTTCCGACTTTGAAACTGCTGTTGCT 1080

Qy 1294 CGTGCACAGCGCGGTGGCTGAGTTCAACCGTCTGCTGTTGCAACCAACATTTGTTTC 1353

Db 1081 CGTGCACAGCGCGGTGGCTGAGTTCAACCGTCTGCTGTTGCAACCAACATTTGTTTC 1140

Qy 1354 TTCGCTGCTGCTGCGGGAAGGAGCTTCACTTCCAGCGCATCCGACCGGATTTATC 1413

Db 1141 TTCGCTGCTGCTGCGGGAAGGAGCTTCACTTCCAGCGCATCCGACCGGATTTATC 1200

Qy 1414 GCGGATCACCACACCTCTTCAGGCTCCACCTGCGGATGATGAGAGGAGCGATCCTG 1473

Db 1201 GCGGATCACCACACCTCTTCAGGCTCCACCTGCGGATGATGAGAGGAGCGATCCTG 1260

Qy 1474 GATTACTTGGCAGATGTCACCGTGAACAGCCTCATGGTGTGGTCCAAAGGATGTTGCA 1533

Db 1261 GATTACTTGGCAGATGTCACCGTGAACAGCCTCATGGTGTGGTCCAAAGGATGTTGCA 1320

Qy 1534 GCACCAATCGATTAAGCTGCCCAACATCAAGGATGTCGCACTGCCACGCGGTTCCCGTGAC 1593

Db 1321 GCTCCTATCGATTAAGCTGCCCAACATCAAGGATGTCGCACTGCCACGCGGTTCCCGTGAC 1380

QY 1594 CGCTGAGCAGCTTGGCCCGCGCG 1620
 |||
 Db 1381 CG-CTGAGCAGCTTGGCCCGCGCG 1406

RESULT 14

SCF11/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SCF11 40104 bp DNA linear BCT 12-MAY-2002
 Streptomyces coelicolor cosmid F11.

AL132662.2 GI:20520901

AL132662.2

3-oxoacyl-(acyl-carrier-protein) synthase; acyl carrier protein;

acyltransferase; alpha-galactosidase; beta-hexosaminidase; cold

shock protein; glucanotransferase precursor; bisindole kinase

protein; hydrolase; oxidoreductase; peptidase; pyruvate

carboxylase; regulatory protein; response regulator; sugar

hydrolase; sugar transporter; transcriptional regulator;

transmembrane protein.

Streptomyces coelicolor A3(2).

Streptomyces coelicolor A3(2).

Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

1 (bases 1 to 40104)

Redenbach, M., Kieser, H.M., Denapate, D., Eichner, A., Cullum, J.,

Kinashi, H. and Hopwood, D.A.

A set of ordered cosmids and a detailed genetic and physical map

for the 8 Mb Streptomyces coelicolor A3(2) chromosome

Mol. Microbiol. 21 (1), 77-96 (1996)

97000351

8843436

2 (bases 1 to 40104)

Seeger, K.J. and Harris, D.

Unpublished

3 (bases 1 to 40104)

Cerdano, A.M., Parkhill, J., Barrell, B.G. and Rajandream, M.A.

Direct Submission

Submitted (26-OCT-1999) Streptomyces coelicolor sequencing project,

Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge

CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.

David A. Hopwood, [3] John Innes Centre, Norwich Research Park,

Colney, Norwich, Norfolk NR4 7UH, UK

On May 9, 2002 this sequence version replaced gi:6137022.

Notes:

Streptomyces coelicolor sequencing at The Sanger Centre is funded

by the BBSRC and Beowulf Genomics

Details of S. coelicolor sequencing at the Sanger Centre are

available on the World Wide Web.

(URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are

numbered using the following system eg SC7B7.01c, SC (S.

coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary

strand).

The more significant matches with motifs in the PROSITE database

are also included but some of these may be fortuitous. The length

in codons is given for each CDS.

Usually the highest scoring match found by fasta -o is given for

CDS which show significant similarity to other CDS in the database.

The position of possible ribosome binding site sequences are given

where these have been used to deduce the initiation codon. Gene

prediction is based on positional base preference in codons using a

specially developed Hidden Markov Model (Krogh et al., Nucleic

Acids Research, 22(22):4768-4778 (1994)) and the FramePlot program

of Bibb et al., Gene 30:157-66 (1984) as implemented at

<http://www.nih.gov/jun/cgi-bin/frameplot.pl>.

CAUTION: We may not have predicted the

correct initiation codon. Where possible we choose an initiation

codon (atg, gtg, ttg or att) which is preceded by an upstream

ribosome binding site sequence (optimally 5-13bp before the

initiation codon). If this cannot be identified we choose the most

upstream initiation codon.

IMPORTANT: This sequence MAY NOT be the entire insert of the

sequenced clone. It may be shorter because we only sequence

overlapping sections once, or longer, because we arrange for a

FEATURES

source

Location/Qualifiers

1. .40104
 /organism="Streptomyces coelicolor A3(2)"
 /strain="A3(2)"
 /db_xref="taxon:100226"
 /clone="cosmid F11"
 252. .255
 274. .1065
 /gene="SCF11.02"
 /note="SCO0522"
 274. .1065
 /gene="SCF11.02"
 /note="SCO0522"
 similar to TR-054197 (EMBL:AJ000671) Streptomyces
 flaviligerus clavulanate-9-aldehyde reductase. 247 aa;
 fasta scores: opt: 501 z-score: 567.7 E(): 3e-24; 39.1%
 identity in 238 aa overlap. Contains Pfam match to entry
 PF00106 adh_short short chain dehydrogenase and proite
 match to entry FS00061 Short-chain
 dehydrogenases/reductases family signature. also similar
 to Streptomyces coelicolor SCF6.15c, 245 aa; fasta scores:
 opt: 818 z-score: 805.2 E(): 0; 56.5% identity in 237 aa
 overlap"
 /codon_start=1
 /transl_table=11
 /product="putative oxidoreductase"
 /protein_id="CAB59579.1"
 /db_xref="GI:6137024"
 /db_xref="SPTREMBL:Q9RK87"
 /translation="MSRTDSAQQRTHRKDIVMTQTSKVVLVTGASSGIGETALRL
 AADHRVFLGARRTERLEKLAIAEDGGTAGYRRLDVTDAADVAFVSAVRWRL
 DVTNNAGVPLSPILSELKVAEDWMLDNNVGRVGHGIAAALPVMAQGGGHVNTAS
 VGAEVVTAAVVCATKFAVRALSEGRLQSAGDIRVSVPGVTSSELADSIDPRA
 REMDKYRSVAVPASAIADAFVSRPAEVDVNEIVVRPAASAQ"
 346. .897
 /genes="SCF11.02"
 /notes="Pfam match to entry PF00106 adh_short, short chain
 dehydrogenase, score 248.20, E-value 1.2e-70"
 751. .837
 /genes="SCF11.02"
 /notes="PS00061 Short-chain dehydrogenases/reductases
 family signature"
 complement(1093).1530
 /genes="SCF11.03c"
 /notes="SCO0523"
 complement(11093).1530
 /genes="SCF11.03c"
 /notes="SCF11.03c, probable transcriptional regulator, len:
 145 aa; similar to SW:ASNC_ECOLI (EMBL:AE000451)
 Escherichia coli, regulatory protein ASNC, 152 aa; fasta
 scores: opt: 222 z-score: 287.1 E(): 1.3e-08; 26.9%
 identity in 145 aa overlap. Contains Pfam match to entry
 PF01037 ASNC_trans_reg, AsnC family"
 /codon_start=1
 /transl_table=11
 /product="putative transcriptional regulator"
 /protein_id="CAB59580.1"
 /db_xref="GI:6137025"
 /db_xref="SPTREMBL:Q9RK86"
 /translation="MDEIDRRLLIALLQDQATRPVAALGCVGLSAGAAHVRKLRQ
 KVRTSTTVDPDEALDRGLVAVLDSTAMWDRADFAVPAVPEIQEAHVIAAGSAALV
 KVTATTERLDQVRLRLVAIDGVSTQATVTVLETFPFRPLDAG"
 complement(1162).1458
 /genes="SCF11.03c"
 /notes="Pfam match to entry PF01037 ASNC_trans_reg, AsnC
 family, score 52.20, E-value 1.1e-11"
 1606. .1610
 1615. .2523
 /genes="SCF11.04"
 /notes="SCO0524"
 1615. .2523

RBS

gene

CDS

misc_feature

misc_feature

gene

CDS

misc_feature

RBS

gene

CDS

```

/gene="SCF11.04"
/note="SCF11.04, possible gluconolactonase, len: 302 aa;
similar to SW:GNL_ZYMO (EMBL:X67189) Zymomonas mobilis
gluconolactonase precursor (EC 3.1.1.17)
D-glucono-delta-lactone lactonohydrolase; fasta scores:
opt: 326 z-score: 392.6 E(): 1.7e-14; 28.7% identity in
258 aa overlap"
/codon_start=1
/transl_table=11
/product="possible gluconolactonase precursor"
/protein_id="CAB59581.1"
/db_xref="GI:6137026"
/db_xref="SPREMBL:Q9RK85"
/transl_table="MSTDGPEYILDDRFTGRGANGDNRLVLYDGRWAEGLYLPA
WRQLVMSDIINDIRLWDEATGSGVFRAPAGHNSGNTLDROGLVTCQGNRRVTRT
EPDRVTVLAEIRDFGLGPNDSVVRSDGTIFSPDFDGTSDYEGHRAESIGACN
VYRIDPVSGOVLRAADGFDGPNVILTPDERRLLFVSDRAARITHAFDIRDGLSDGK
VFAEGRGDVAHDIRFDEDEGLRWAAALHDGVCHYCDPDGTLIGRLVPEPVSNAFGGP
KNKRLFITATTSLSYLSVMSVTGAPRL"
2654. .3469
/gene="SCF11.05"
/note="SCF11.05"
/db_xref="SC00525"
2654. .3469
/gene="SCF11.05"
/note="SCF11.05, hypothetical protein, len: 271 aa;
similar to various hypothetical proteins, e.g. TR:CAB45341
(EMBL:AL079345) Streptomyces coelicolor SC868.05C
hypothetical 33.1 kD protein, 300 aa; fasta scores: opt:
808 z-score: 960.4 E(): 0; 53.2% identity in 250 aa
overlap"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein SCF11.05"
/protein_id="CAB59582.1"
/db_xref="GI:6137027"
/db_xref="SPREMBL:Q9RK84"
/transl_table="MSDPSTPGHTAQOKIDTSVPHSARIWNVLGKKNYPVDEOAG
DAYTAVFPGTVVARSARAFRLRNITHVAEGRFELDITGLTADNTHVAQRIA
PESRIVVDVMTVLHARALLYSTPEGATSYDSVDLPDPTVLEAAATLDFDRPIA
LILSNILGHPIDYDOARSVAARLVDALPSGSYLSINDSGRIDPEPDRADGYNESA
APVILRTVDOITAFDFGLLELVEPGVVPVTRWDATEDFGLIGEHGGLRKP"
3493. .3497
3507. .4340
/gene="SCF11.06"
/note="SCF11.06"
/note="SC00526"
3507. .4340
/gene="SCF11.06"
/note="SCF11.06, possible oxidoreductase, len: 277 aa;
similar to TR:O53622 (EMBL:AL021428) Mycobacterium
tuberculosis putative oxidoreductase MT030.21c, 276 aa;
fasta scores: opt: 276 z-score: 329.4 E(): 5.6e-11; 24.7%
identity in 267 aa overlap and to SW:PRXC_PSEFL
(EMBL:AF031153) Pseudomonas fluorescens non-heme
chloroperoxidase (EC 1.11.1.10) (chloride peroxidase)
(CPO-F) (chloroperoxidase F) Cpo or CpoF; fasta scores:
opt: 233 z-score: 279.3 E(): 3.5e-08; 26.1% identity in 283
aa overlap. Contains Pfam match to entry PF00561
abhydrolase, alpha/beta hydrolase fold"
/codon_start=1
/transl_table=11
/product="possible oxidoreductase"
/protein_id="CAB59583.1"

```

```

Query Match      38 9%; Score 1351; DB 1; Length 40104;
Best Local Similarity 63.5%; Pred. No. 1.7e-280;
Matches 2154; Conservative 0; Mismatches 1215; Indels 24; Gaps 5;

Qy 85 TTCAAAAGATCTTGGTAGCAAAACCGCGGAAATCGCGTCCGTCCTTCCTCGTCAGCA 144
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32519 TTCCGCAAGTGCTGGTCCCAACCGCGGTGAGATCGGATCCGTGCGTTTCGGGGGGC 32460
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 145 CTCGAAACCGGTGAGCCACCGGTAGCTATTATCCCCCGTGAAGATCGGGGATCATTCAC 204
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 32459 TACGAGCTCGCGCGCCACCGCTCGTCCGCGACGAGGACCGCAATTCTCTCTGAC 32400
Qy 205 CGCTCTTTTGTCTCTGAAGCTGTCCGCAATGTGTAAGGCTCACCAGTCAAGCGGTAC 264
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32399 CGGCTCAAGCGCCGACGAGGCTACGAGATCGGGGAGCAGGGGCATCCGTCGCGGTAC 32340
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 265 CTGGACATCGATGAATATTATCGGTGAGCTAAATAAGTTAAAGCAGATGCTATTATCCCG 324
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32339 CTCTCCGTGGAGGAGATCGTCCGCGCCGCCGCGTTCGGGGGCGACGCGCTACTCCG 32280
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 325 GGATATGGCTTCCTGTCTGAAAATGCCAGCTTGCCCGCGAGTGCAGGAAACCGGAT 384
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32279 GGCTACGGCTTCCTGTCCGAGAACCCCGAACTCGCCCGCGCTCGAGAGGCGCGGATC 32220
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 385 ACTTTTATTGGCCCAACCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCTGTCGGTA 444
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32219 ACCTTCGTGCTCCAGCGCCCGGATCTTGAACTGACCGGCAACAAGGACGCGGCGTG 32160
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 445 ACCGCGCGAAGAGGCTGTGTCAGATTTTGGCGGAATCCACCCCGAGCAAAAACATC 504
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32159 GCGCGCGCCGCGAGGCGCGAGTACCGGTCTCTCGGCTCTCGGCGCTCTCCACGCGTG 32100
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 505 GATGACATCGTTAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTAAGGCGAGTTGCC 564
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32099 GACGAATCTGTACGCGCGCGCGAGACGTGCGCTTCCCGGTTCGTCAGAGCGGTGCG 32040
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 565 GGTGTTGCGGAGCGGTATGCGCTTTTGTTCCTCACTGATGAGCTCCGCAAAATTTGCA 624
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32039 GCGGCGCGGCGCGGATGCGCGCGTCAAGGAAACCCGCCAGCTTCGCGGAGCCATC 31980
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 625 ACAGAAGCATCTCGTGAAGCTGAAGCGCATTCGCGCAAGGTTTCGGTATATGTGCAAGT 684
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31979 GAGCGCGCTCCGCGAGGCGCGCTTCGCGCTTCGCGGACTCCACCGCTCTCTCGAGAAG 31920
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 685 GCTGTGATTAACCCAGCACATTAAGTGCAGATCCTTGGCGATCCACTGGAGAGTT 744
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31919 GCGGTCTCGAAACCCGCCACATCGAGGTGAGATCTCGCGCGAGGCGGAGCGTC 31860
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 745 GTACACCTTTATGAACGTACTGTCTCACTGAGCGTCTGTCACCAAAAAGTTTGTGCAAA 804
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31859 ATCCACCTCTTCGAGCGGACTGCTCGTGCAGCGCCGCCACAGAAAGTTCATCGAGCTG 31800
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 805 GCGCGACGACAGCATTTGGATCCAGAACTGCGGTATCGCATTTGTTCGGATGCAAGTAAAG 864
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31799 GCGCGCGCGCAACCTCGACCCGCGCTCGCGGAGCGGATCTGCGCGCAGCGCGTGAAC 31740
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 865 TTCTGCGCTCCATTGTTTACCAGGCGCGGACCGTGGATTCCTTGGTCGATGAAAAG 924
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31739 TTGCGCGCGGAGATCGGCTACCGCAACGCGGCAACCGTTCAGGTTCGTCGACCGGGAC 31680
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 925 GGCACACCGTTTTTCATCGAAATGAACCCACGTTATCCAGGTTGAGCACACCGTGACTGAA 984
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31679 GGCACACCGTTTCATCGAGATGAACCCCGGATCCAGGTTCGAGCACACGTCACCGAG 31620
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 985 GAAGTACCGAGGTGACCTGGTGAAGGCGCAGATCGCTTGGCTGTGTTGTCGAACCTTG 1044
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31619 GAGTACCGGAGTGCAGCTGGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 31560
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1045 AAGGAATGGGTCTGACCCCAAGATTAAGATCAAGACCCACCGTGCAGCACTCGAGTGGCG 1104
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31559 GCGGACCTCGGATCGCCCGAGGAGAACATCACCTGCGCGGTGCGGCTCCAGTCCAGTGGCG 31500
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1105 ATACACCGAAGATCCAAACAAACGCGCTTCGCGCCAGATACCGGAATATCACCGCGTAC 1164
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31499 ATACACCGAGAGACCGCGCAACGCGTTTCGCGCGGACACCGGGAGATCAGCGCTAC 31440
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1165 CGCTACCGAGCGAGCTGCGTTCGTTGA---CGTGCAGCTCAGCTCGTGGGCGCAA 1221
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31439 GGTTCGCGGCGGCTCCGCGATCCCGGCTCGAGCGGCGTACCACCCACGCGGTACGAG 31380
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1222 ATACCGGACACTTTGATTCATGCTGCTGTAATAATGACTTGCCTGGTTCGATTTGAA 1281
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31379 ATACGCGCACTTCGATCGATGCTGTAAGCTCTCTCGCGGGGACGGGACTTCACC 31320
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Db 29180 GAGGCGGCGATCTTCTCGTCAACTCGCCTGA 29148

RESULT 15

LOCUS AF262949

DEFINITION Mycobacterium smegmatis hypotheical integral membrane protein gene, partial cds; pyruvate carboxylase (pyc) gene, complete cds; and hypotheical methyltransferase gene, partial cds.

ACCESSION AF262949

VERSION AF262949.1

KEYWORDS GI:11118647

SOURCE Mycobacterium smegmatis.

ORGANISM Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

REFERENCE 1 (bases 1 to 4218)

AUTHORS Mukhopadhyay,B. and Purwantini,E.

TITLE Pyruvate carboxylase from Mycobacterium smegmatis: stabilization, rapid purification, molecular and biochemical characterization and regulation of the cellular level

JOURNAL Biochim. Biophys. Acta 1475 (3), 191-206 (2000)

MEDLINE 20374587

PUBMED 10913817

REFERENCE 2 (bases 1 to 4218)

AUTHORS Mukhopadhyay,B. and Purwantini,E.

TITLE Direct Submission

JOURNAL Submitted (02-MAY-2000) Microbiology, University of Illinois, 601 South Goodwin Avenue, Urbana, IL 61801, USA

FEATURES

source

1. 4218

/organism="Mycobacterium smegmatis"

/db_xref="taxon:1772"

<1_564

/note="ORF1; similar to putative integral membrane protein of Mycobacterium leprae encoded by GenBank Accession Number AL023635 and to hypotheical protein Rv2968c of Mycobacterium tuberculosis strain H37Rv encoded by GenBank Accession Number Z83018"

/codon_start=1

/transl_table=11

/product="hypotheical integral membrane protein"

/protein_id="AAG30410.1"

/db_xref="GI:11118648"

/translation="PSAVWVLIAGVLGLAASLTITVEKIELLINPDVPSCSINPVLSCGSMWVWQSLGFPNPLIGIVAFSVLVTVGLAVAGVRLPRWYAGLATGTALGAVFVHLIFQSLIRIGALCPYCMVWVAVTIPLAVVTVAVIAURAGADADAGPENVHQRWSLVLMFTALVLLILVRFWEYWSLL"

600..3983

/gene="pyc"

600..3983

/gene="pyc"

/EC_number="6.4.1.1"

/function="ATP-dependent pyruvate carboxylase"

/note="purified native enzyme shows ATP-dependent pyruvate carboxylase activity in vitro; NH2-terminus of mature pyc protein confirmed by Edman degradation"

/codon_start=1

/transl_table=11

/product="pyruvate carboxylase"

/protein_id="AAG30411.1"

/db_xref="GI:11118649"

/translation="MISKVLVNRGETAIPRAAYEMGIATVAVVPEYEDNSLHLKADSSYQIGEVCHPVRAVLSVDEIIRVAKHSGADAVPGYGLFSENFDLAKAEAGITFVPSAEVLQTNKAKIAAARAAGLPVLSSESPSSVDLMMAAADWEFFLPFVXAVSGGGRMRVRESRLAEIAEASREAESAFGDASVYLEQAVLNPRHIEVQILADGAGNWHLPFERDCSVORHQVKEVLPAPNLSDELROQICADAVAFQIIGYSCAGTVFLLDERGHVFIENCPNTOVEHTTIDVDLVSSOLRIAAGETLADLIGLSODRLVVRGAACQCHITTEVPANGRPDTRITAYRSPGGAGIRLDGNTNLGAISLHAFDSMLVLLTCRGDFSAASARRALAEFRINGRSTNIPFIQAVIDDDPFRAGRVTTFSFIDDRPHLLTGRSPADRGTRILNLYLADITVKNKPHGERPSTVYPQDKLPFLDLQAPPAGSKQRILVE

LGPEGAGWLRBSKAVGVTDTTFRDAHQSLATRVTTGLLMVAVPVYARSPQLLSLE
CMGATYDVALRFLKEDPWERLAALRESVFNICQMLLRGNTVGTYPPELVTSAYV
EEAATGIDIFRIFDALNNVESMRPAIDAVRETGSTIAEAMCYTGDLSDAENLYTL
DYTLKAEQIVERAGAHVLAIKOMAGLRLPAHAHTLVASLRSPDLPVHVHTDTPGQ
LATYLAAGADAVDASAPMAGTTSQALSSIVAAAHQTDVDTGLDRAVCDLBPY
WEAVRVYAPFSSGLPGPTGRVYVTHEIRGGQSLNRQQAIALGLGDRFEERAA
DRVLGKVTTPSSKVVGBLALALVAGAGITAEFAEDPAKYDIPDSVIGLSELGDP
PVGWPLPKTKALQGGPARPVKLTADDEALLAQPGKQAALNRLFLPGPTAEFPA
HRETYDTSLSANQFFYGLRYGEEHVRQLERGVLLIGLEAISADERGMRTVMCI
NGQLRVLVRDSIASVEPAEKARNADHIAAPFAGVTVTVGVAEGSDVADGQTAT
IEAMKMEAAITAPKAGTAVARVAATAQVEGGDLLVVS"

misc_feature 600..2159

/gene="pyc"

/note="Region: PYCA or ATP-binding domain"

misc_feature 1455..1478

/gene="pyc"

/note="Region: PS00867 carbamoyl-phosphate synthase subdomain signature 2"

misc_feature 2160..3980

/gene="pyc"

/note="Region: PYCB or the biotin carrying domain"

misc_feature 2679..2720

/gene="pyc"

/note="Region: PS00165 serine/threonine dehydratases pyridoxal-phosphate attachment site"

misc_feature 3867..3884

/gene="pyc"

/note="Region: PS00188 biotin-requiring enzymes attachment site"

CDS 3980..54218

/note="ORF2; contains N-6 adenine-specific DNA methylase signature, PS00092; similar to hypotheical protein MLCB1243.09 of Mycobacterium leprae encoded by GenBank Accession Number AL023635 and to hypotheical protein Rv2966c of Mycobacterium tuberculosis strain H37Rv encoded by GenBank Accession Number Z83018"

/codon_start=1

/transl_table=11

/product="hypotheical methyltransferase"

/protein_id="AAG30412.1"

/db_xref="GI:11118650"

/translation="WTRIIAGFLGRRRIEVPKSGTRPTSDRVREAVFNALALDFAGLAVLDLYAGSGALGLEAISRGASSLUFVSDQRAA"

BASE COUNT 623 a 1433 c 1479 g 683 t

ORIGIN

Query Match 38.3%; Score 1329.8; DB 1; Length 4218;

Best Local Similarity 62.8%; Pred. No. 6.9e-276;

Matches 2123; Conservative 0; Mismatches 1257; Indels 9; Gaps 3;

Qy 86 TCAAAAAGATCTTGGTAGCAAAACCGCGGCAAAATCGCGGTCCGTGCTTTCCGTGCAGCAC 145

Db 604 TCTCCAGGTGCTCGTGCCTCAACCGCGGCGAATCGCGCATTCGGTCTGGGT 663

Qy 146 TCGAAACCGGTGCAGCACCGGTAGCTATTTACCCCGGTGAAGATCGGGATCATTCAC 205

Db 664 ACAGATGGGCATCGCCACCGTGGCGGTGTATCGTACGAGGACCGGAATTCGCTCCATC 723

Qy 206 GCTCTTTTGTCTTGAAGCTGTCCGATTTGCTAGGCTCACCAGTCAAGCGGTACC 265

Db 724 GGCTCAAGCGCGACGATCATATCAGATCGCGAGGTGGGTGTCATCCGTCCGCGGTATC 783

Qy 266 TGGACATCGATGAAATTTATCGGTGCAGCTAAAAAGTTAAAGCAGATGCTATTATTACCGG 325

Db 784 TGTCCGTGCAGAGATCATTCGGGTGCCAAGCATTCGGGCGCGACGCGGTGTACCGG 843

Qy 326 GATATGGCTTCTGTCTGAAAAATGCCAGTTCGCCCGAGTGGCGGGAACAAACGGCATTA 385

Db 844 GCTACGGCTTCTGTCTGGAGAACCCCGATCTGCGGCGCAAGTGGCGGCGGGTATCA 903

Qy 386 CTTTTATTGGCCCAACCCAGAGAGTTCTTGATCTCACCGGTGATAAGTCTCGTGGCGGTAA 445

Db 904 CGTTCGTGGGACCGTCCGCGCGAGGTGCTGAGCTACGGGTAAACAGGCAACGCGCGATCG 963


```
Db 3124 TCGGCGACCGGTTGAGAGATCGAGGCAATTACGCTGCGCGGACCCGGGTTCGGGAC 3183
Qy 2666 GCCCAACCAAGGTCAACCCATCTCTCAAGTTGTTGGGAGCTCGCACTCCACCTGGTTG 2725
Db 3184 GGCTCGTGAAGGTGACCCCGCTCGTGAAGTGGTCGGGACCTGGCGCTCGGTG 3243
Qy 2726 GTGGGGTGTAGATCCAGACACTTTGTGCAGACCCACAAAGTACGACATCCAGACT 2785
Db 3244 GTGGGGGATCACCGCCGAGGATTCGCCGAGGATCCCGGAAGTACGACATCCCGACA 3303
Qy 2786 CTGTCATCGCTTCTCGCGCGGAGCTTGTAACTCCAGGTGGCTGGCCAGAACCA 2845
Db 3304 GCGTATCGGCTTCTCGCGGCTGAATCGGGATCCCGCGGATGGCCGGAACCGT 3363
Qy 2846 TCGCACCCGCGCACTGGAAGCCGCTCCGAAGCAAGCACCTCTGACGGAAAGTTCCGT 2905
Db 3364 TCGCACCAAGCGCTCCAGGCGCGGACCGGCCCGG--CCGCTCGAAGCTGACCG 3420
Qy 2906 AGGAAGACGAGCGCACTCGACGCTGATGATTCAGGAACGTGCGCAACAGCCTCAAC 2965
Db 3421 CCGACGACGAGCGCTGTGCCCCAGCCCGGCCCAAG---CGGACGGCCGCTTGAACC 3477
Qy 2966 GCCTGCTGTCCCGAAGCAACCGAGAGTTCCTCGAGCACGCTCGCGCTTCGGCAACA 3025
Db 3478 GCCTGCTTTTCCCGGGCCACCGCGAGTTCGAGGCGCACCGGAACCTACGGCGACA 3537
Qy 3026 CCTCTGCGCTGGATGATCGTGAATTTCTTACGGACTGCTCGAGGGCCGCGAGACTTTGA 3085
Db 3538 CCTCATCCTCAGCGGAACCAAGTTCTTACGGGCTCGCTACGGCGAGGACCCCGG 3597
Qy 3086 TCGGCTGCCAGATGTGGCGACCCCACTGCTTTGTTGCTGGCTGGATCTCTAGCCAG 3145
Db 3598 TCGAACTCGAACGCTG---GCTGGAACTGCTGATCGGGCTTGAAGCGATCTCGGAGCGG 3654
Qy 3146 ACGATAAGGGTATGCGCAATGTTGTGGCCACGTCACCGCGAGATCGCCCAATGCGTG 3205
Db 3655 ACGAGCGGGCATGCGCACCGCTGATGTCATCAACGGTCAGCTGCGCCCGGTTCTCG 3714
Qy 3206 TCGGTGACCGCTCGGTTGAGTCTGTACCGCAACCGCAAAAGGAGATTCCTCCAACA 3265
Db 3715 TCGCGACCGGAGCATCGCCAGGAGTGCCCGCGCGAAGGCGGACCGCAACATG 3774
Qy 3266 AGGGCCATGTTGTGCAACCATTCGCTGGTGTGTACTGTGACTGTTGCTGAAGGTGATG 3325
Db 3775 CCGACCAATCGCCGCGCCTTTCGCCGCTGTGGTGACCGTGGTGTGCGAAGGTGACT 3834
Qy 3326 AGGTCAAGGCTGGAGATGCACTCGCAATCATCGAGGCTATGAAGATGGAAGCAACATCA 3385
Db 3835 CGGTGGACCGGGACAACCATCGGACGATCGAGGCGATGAAGATGGAGGCCGCCATCA 3894
Qy 3386 CTGCTTCTGTTGACGGCAAGATTGAACGCTTGTGGTTCTCTGCTGCAACGAAGGTGAAG 3445
Db 3895 CCGGCCCAAGCGAGGACCGTTCGGCGCGTTCGGGTGCGCGGACCGCCAGGTGAGG 3954
Qy 3446 GTGGCGACTTCATCGTCTGTTCTTA 3474
Db 3955 GCGCGATCTGCTGTGTGTGCTGACTGA 3983
```

Search completed: April 5, 2003, 10:09:08
Job time : 10481 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2003, 04:30:25 ; Search time 762 Seconds
(without alignments)
10266.984 Million cell updates/sec

Title: US-09-974-973-1
Perfect score: 3474
Sequence: 1 gtagctgtatcaccttgg.....tgatgctgctgttctctaa 3474

Scoring table:
IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3474	100.0	3474	24	ABK52832
2	3398.8	97.8	3621	21	AAA47533
3	3398.8	97.8	3621	22	AAE32165
4	3398.8	97.8	349980	22	AAH68526
5	3397.2	97.8	3728	20	AAZ4102
6	3344.8	96.3	3420	22	AAH65730
7	2256.2	64.9	4013	22	AAE7437
8	1564.8	45.0	1719	22	AAE71419
9	1351.8	38.9	1406	22	AAE71420
					Corynebacterium gl
					Sequence encoding
					Corynebacterium gl
					C glutamicum codin
					C. glutamicum pyru
					C. glutamicum codin
					Corynebacterium th
					Corynebacterium gl
					Corynebacterium gl

10	1020.2	29.4	1083	22	AAF71417	Corynebacterium gl
11	1020.2	29.4	1083	22	AAF71418	Corynebacterium gl
12	883	25.4	939	22	AAF71415	Corynebacterium gl
13	883	25.4	939	22	AAF71416	Corynebacterium gl
14	681	19.6	3945	24	ABK63659	Rat sequence diffe
15	674.2	19.4	4017	24	ABN96532	Gene #3030 used to
16	665.2	19.1	4138	23	ABL19577	Drosophila melanog
17	665.2	19.1	4308	23	ABL19575	Drosophila melanog
18	665.2	19.1	4369	23	ABL202789	Drosophila melanog
19	620.2	17.9	4034	23	ABL20985	Drosophila melanog
20	606	17.4	4152	24	ABO70879	Listeria monocytog
21	565.8	16.3	3011208	24	ABO69245	Listeria innocua D
22	564	16.2	3462	22	AA501509	Bacillus subtilis
23	502.8	14.5	3429	23	AA553072	Enterococcus faeca
24	488.2	14.1	20072	20	AAK13026	Enterococcus faeca
25	486.8	14.0	3441	23	AA551831	Staphylococcus aur
26	476.6	13.7	3456	22	AAH53527	S. epidermidis ope
27	476.6	13.7	3465	24	ABN91128	Staphylococcus epi
28	464.4	13.4	5030	18	AAV74635	Staphylococcus aur
29	428.8	12.3	3222	23	AA554626	Staphylococcus aur
30	345.4	9.9	3432	22	AAH54232	S. epidermidis gen
31	308.2	8.9	7489	23	ABL20984	Drosophila melanog
32	308.2	8.9	7703	23	ABL19576	Drosophila melanog
33	308.2	8.9	12028	23	ABL02788	Drosophila melanog
34	308.2	8.9	12039	23	ABL19574	Drosophila melanog
35	305	8.8	1362	17	AAAT3071	Synechococcus biot
36	305	8.8	1362	19	AAV33406	Synechococcus acet
37	303.4	8.7	1362	15	AAQ45330	fabg gene encoding
38	302.2	8.7	1350	23	AA552596	E. coli DNA for ce
39	302.2	8.7	3077	16	AAQ79920	E. coli acb and a
40	302.2	8.7	14000	22	AA546232	DNA encoding novel
41	297	8.5	3065	15	AAQ45328	fabg gene encoding
42	297	8.5	3065	17	AAAT3070	Anabaena biotin ca
43	297	8.5	3065	19	AAV33405	Anabaena acetyl-Co
44	286	8.2	2238	24	ABK72711	Bacillus lichenifo
45	283.8	8.2	1350	23	AA556150	Salmonella typhi D

ALIGNMENTS

RESULT 1
ID ABK52832 standard; DNA; 3474 BP.
XX
AC ABK52832;
XX
DT 27-AUG-2002 (first entry)
XX
DE Corynebacterium glutamicum feedback-resistant pyruvate carboxylase gene.
XX
KW Feedback-resistant; pyruvate carboxylase enzyme; gene; ds;
KW aspartic acid feedback inhibition resistant.
XX
OS Corynebacterium glutamicum.
XX
FH Key Location/Qualifiers
FT CDS 1..3474
FT /tag= a
FT /product= "Corynebacterium glutamicum feedback-resistant
pyruvate carboxylase enzyme"
FT /transl_except= (pos:1..3; aa:Met)

WO200231158-A2.

18-APR-2002.

12-OCT-2001; 2001WO-US31893.

13-OCT-2000; 2000US-239913P.

(ARCH) ARCHER-DANIELS MIDLAND CO.

PI Hanke PD:
XX WPI; 2002-463267/49.
DR P-PSDB; AAU98050.
XX
PT Novel mutated, feedback resistant pyruvate carboxylase enzyme
PT polypeptide, useful for producing amino acids e.g. L-lysine,
PT L-threonine, L-glycine, L-glutamic acid, L-proline and L-methionine and
PT L-isoleucine -
XX
PS Claim 3; Fig 1; 42pp; English.
XX
CC The present invention relates to a new mutated, feedback-resistant
CC pyruvate carboxylase enzyme. The invention is useful for producing an
CC amino acid (e.g. L-lys, L-thr, L-Met, L-Ile, L-Glu, L-Arg and L-Pro),
CC by culturing a host cell in a suitable media and separating the amino
CC acid from the medium. The vector of the invention is useful for
CC replacement of a wild-type pyruvate carboxylase gene, with a feedback
CC resistant pyruvate carboxylase gene, in a Corynebacterium glutamicum, by
CC replacing a genomic copy of the wild-type pyruvate carboxylase gene with
CC a selectable marker gene through homologous recombination to form a first
CC recombination strain, and replacing the selectable marker gene in the
CC first recombinant strain, with feedback resistant pyruvate carboxylase
CC gene through homologous recombination to form a second recombinant
CC strain, where the homologous recombination in the above steps, occurs
CC between the host cell and the vector. The feedback-resistant pyruvate
CC carboxylase enzyme is resistant to feedback inhibition from aspartic
CC acid. The present nucleic acid sequence encodes the feedback-resistant
CC pyruvate carboxylase enzyme of the invention.
XX
SQ Sequence 3474 BP; 749 A; 1013 C; 950 G; 762 T; 0 other;

Query Match 100.0%; Score 3474; DB 24; Length 3474;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCACTGCTATCAACCTCGCGCTCTCTGTGAAGGAATAATCTACTGTCGACT 60
DB 1 GTGACTGCTATCAACCTCGCGCTCTCTGTGAAGGAATAATCTACTGTCGACT 60
QY 61 CACACATCTTCAACGGTTCAGGATTCAAAAGATCTTGTAGCAAAACCGCGGGAATC 120
DB 61 CACACATCTTCAACGGTTCAGGATTCAAAAGATCTTGTAGCAAAACCGCGGGAATC 120
QY 121 GCGGTCCGTCTTCCGTGAGGACTCGAAACCGGTGAGCCAGGTAGCTATTTACCCC 180
DB 121 GCGGTCCGTCTTCCGTGAGGACTCGAAACCGGTGAGCCAGGTAGCTATTTACCCC 180
QY 181 CGTGAAGATCGGGGATCATTCACCGCTCTTTTGTCTGAAGCTGTCCGCTTGTACT 240
DB 181 CGTGAAGATCGGGGATCATTCACCGCTCTTTTGTCTGAAGCTGTCCGCTTGTACT 240
QY 241 GAAGGCTCACCAGTCAAGCGTACCTGGACATCGATGAAATTTATCGGTGAGCTTAAAA 300
DB 241 GAAGGCTCACCAGTCAAGCGTACCTGGACATCGATGAAATTTATCGGTGAGCTTAAAA 300
QY 301 GTTAAAGCAGATCTATTTACCGGGATATGCTTCTGCTGAAATGCCAGCTTGGC 360
DB 301 GTTAAAGCAGATCTATTTACCGGGATATGCTTCTGCTGAAATGCCAGCTTGGC 360
QY 361 CGCGAGTGGCGGAAACCGGATTAATTTATGGCCAAACCGGAGGTTCTTGATCTC 420
DB 361 CGCGAGTGGCGGAAACCGGATTAATTTATGGCCAAACCGGAGGTTCTTGATCTC 420
QY 421 ACCGGTATTAAGTCTCGTGGGTAAACCGGCGGAAGAGGCTGTCTGCGAGTTTGGCG 480
DB 421 ACCGGTATTAAGTCTCGTGGGTAAACCGGCGGAAGAGGCTGTCTGCGAGTTTGGCG 480
QY 481 GATTCACCCCGAGCAAAAACATCGATGACATCGTTAAAAAGCGTGAAGGCCAGACTTAC 540
DB 481 GATTCACCCCGAGCAAAAACATCGATGACATCGTTAAAAAGCGTGAAGGCCAGACTTAC 540
QY 541 CCCATCTTTGTAAAGGACGTTGCGGTGGTGGCGGACCGCGTATGCGCTTTGTTCTTCA 600

DB 541 CCCATCTTTGTAAAGGACGTTGCGGTGGTGGCGGACCGCGTATGCGCTTTGTTCTTCA 600
QY 601 CTTGATGAGCTCCGCAAAATTTGCAACAGAGCATCTCGTGAAGCTGAAGCGGATTCGGC 660
DB 601 CTTGATGAGCTCCGCAAAATTTGCAACAGAGCATCTCGTGAAGCTGAAGCGGATTCGGC 660
QY 661 GACGGTTCCGTATATGTCGAACGCTGTGATTAACCCCGAGCAGCATTTGAAGTGCAGATC 720
DB 661 GACGGTTCCGTATATGTCGAACGCTGTGATTAACCCCGAGCAGCATTTGAAGTGCAGATC 720
QY 721 CTTGCGGATCGACCTGAGAGAGTTGTACACCTTTATGAACGTGACTGCTCAGTCAGCGT 780
DB 721 CTTGCGGATCGACCTGAGAGAGTTGTACACCTTTATGAACGTGACTGCTCAGTCAGCGT 780
QY 781 CGTCACAAAAAGTTGTGAAAATTTGCGCGCAGCAGCATTTGGATCAGAACTGCGGTAT 840
DB 781 CGTCACAAAAAGTTGTGAAAATTTGCGCGCAGCAGCATTTGGATCAGAACTGCGGTAT 840
QY 841 CGCAATTTGCGGATGCAAGTAAAGTTCTGCGCTCCATTTACACAGGCGCGGGAACC 900
DB 841 CGCAATTTGCGGATGCAAGTAAAGTTCTGCGCTCCATTTACACAGGCGCGGGAACC 900
QY 901 GTGGAATCTTCTGTCGATGAAAAGGGCAACCAACCTTTTCATCGAAATGAAACCCACGTATC 960
DB 901 GTGGAATCTTCTGTCGATGAAAAGGGCAACCAACCTTTTCATCGAAATGAAACCCACGTATC 960
QY 961 CAGGTTGAGCACACCGTGCATGAAAGAGTCAACCGAGGTGACCTGTGTGAAGGCGCAGATG 1020
DB 961 CAGGTTGAGCACACCGTGCATGAAAGAGTCAACCGAGGTGACCTGTGTGAAGGCGCAGATG 1020
QY 1021 CGCTTGGCTGCTGGTGCACACCTTTGAAGGAATTTGGGTCTGACCCAAAGATAGATCAAGACC 1080
DB 1021 CGCTTGGCTGCTGGTGCACACCTTTGAAGGAATTTGGGTCTGACCCAAAGATAGATCAAGACC 1080
QY 1081 CAGGTTGAGCACCTCGACTCGCGGTGCGCATCACACGGAAGATCCAAACACGCTTCCGCCCA 1140
DB 1081 CAGGTTGAGCACCTCGACTCGCGGTGCGCATCACACGGAAGATCCAAACACGCTTCCGCCCA 1140
QY 1141 GATACCGGAATCTACACCGGTACCGCTCACAGCGGAGCTGGCTTGTGTGAGCGT 1200
DB 1141 GATACCGGAATCTACACCGGTACCGCTCACAGCGGAGCTGGCTTGTGTGAGCGT 1200
QY 1201 GCAGTCAAGTCCGTGGGCAAAATCAACGACATCTTGAATTCATGCTGTGTGAAATGACC 1260
DB 1201 GCAGTCAAGTCCGTGGGCAAAATCAACGACATCTTGAATTCATGCTGTGTGAAATGACC 1260
QY 1261 TGCCTGTGTTCCGACTTTGAAAATGCTGCTGCTGACAGCGCGCTTGGCTGAGTTTC 1320
DB 1261 TGCCTGTGTTCCGACTTTGAAAATGCTGCTGCTGACAGCGCGCTTGGCTGAGTTTC 1320
QY 1321 ACCGTGCTGTGTTGCAACCAACATTTGGTTTCTTGGCTGCTGCTGCGGGAAGAGGAC 1380
DB 1321 ACCGTGCTGTGTTGCAACCAACATTTGGTTTCTTGGCTGCTGCTGCGGGAAGAGGAC 1380
QY 1381 TTCACCTTCAAGCGCATCGCCACCGGATTTATCGGCGATCACCCACACCTCTTCAGGCT 1440
DB 1381 TTCACCTTCAAGCGCATCGCCACCGGATTTATCGGCGATCACCCACACCTCTTCAGGCT 1440
QY 1441 CCACCTGCGGATGATGAGCAGGAGCGCATCTCGGATTAATTTGGCAGATGTCACCGTGAAC 1500
DB 1441 CCACCTGCGGATGATGAGCAGGAGCGCATCTCGGATTAATTTGGCAGATGTCACCGTGAAC 1500
QY 1501 AAGCTCATGCTGCTGGTCCAAAGGATTTGACAGCACAATCGATGAAGTGCAGCAATC 1560
DB 1501 AAGCTCATGCTGCTGGTCCAAAGGATTTGACAGCACAATCGATGAAGTGCAGCAATC 1560
QY 1561 AAGGATCTCCCATGCGCAGCGGTTCCCGTGAACCGCTGAAAGAGCTTGGCCAGCGCG 1620
DB 1561 AAGGATCTCCCATGCGCAGCGGTTCCCGTGAACCGCTGAAAGAGCTTGGCCAGCGCG 1620
QY 1621 TTTGCTCGTGTATCTCCGTGAGCAGGACGCTGCGAGTTTATGATACCACTTTCGCGAT 1680

Db 1621 TTTGCTCGTGATCTCCGTGACGAGACCGCACTGGCAGTTACTGATACCACTTCGCGAT 1680
Qy 1681 GCACACAGTCTTTGCTTGGCGACCGAGTCCGCTCATTCGCACTGAAGCTTGGCGCAGAG 1740
Db 1681 GCACACAGTCTTTGCTTGGCGACCGAGTCCGCTCATTCGCACTGAAGCTTGGCGCAGAG 1740
Qy 1741 GCGTCCGAAAGCTGACCTCTGAGCTTTTGTTCCTGGAGGCTGGGGCGGCGACCTAC 1800
Db 1741 GCGTCCGAAAGCTGACCTCTGAGCTTTTGTTCCTGGAGGCTGGGGCGGCGACCTAC 1800
Qy 1801 GATGTGGCGATGCGCTTTCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGGAG 1860
Db 1801 GATGTGGCGATGCGCTTTCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGGAG 1860
Qy 1861 GCGATGCCGAATGAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACAC 1920
Db 1861 GCGATGCCGAATGAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACAC 1920
Qy 1921 CCGTACCAGACTCCGCTCGCGCGCTTTGTTAAGGAAGCTGCCAGCTCCGCGGTGGAC 1980
Db 1921 CCGTACCAGACTCCGCTCGCGCGCTTTGTTAAGGAAGCTGCCAGCTCCGCGGTGGAC 1980
Qy 1981 ATCTTCCGCATCTTCGACCGCGCTTAAAGACGCTCTCCAGATGCGTCCAGCAATCGACGA 2040
Db 1981 ATCTTCCGCATCTTCGACCGCGCTTAAAGACGCTCTCCAGATGCGTCCAGCAATCGACGA 2040
Qy 2041 GTCTCGGAGACCAACACCGCGGTAGCCAGGTGCTATGGCTTATTCGTGTGATCTCTCT 2100
Db 2041 GTCTCGGAGACCAACACCGCGGTAGCCAGGTGCTATGGCTTATTCGTGTGATCTCTCT 2100
Qy 2101 GATCCAAATGAAGACTCTACACCTGGATTACTACCTAAGATGGCAGGAGATCGTC 2160
Db 2101 GATCCAAATGAAGACTCTACACCTGGATTACTACCTAAGATGGCAGGAGATCGTC 2160
Qy 2161 AAGTCTGCGCTCACATTTCTGCGCAATTAAGGATATGCTGTCTGCTCGCCAGCTGCG 2220
Db 2161 AAGTCTGCGCTCACATTTCTGCGCAATTAAGGATATGCTGTCTGCTCGCCAGCTGCG 2220
Qy 2221 GTAAACCAAGCTGGTCAACCGACTCGCGCGTGAATTCGATCTGCCAGTGCACGCAACC 2280
Db 2221 GTAAACCAAGCTGGTCAACCGACTCGCGCGTGAATTCGATCTGCCAGTGCACGCAACC 2280
Qy 2281 CACGACACTCGGGTGGCAGTGGCTACCTACTTTGCTGAGCTCAAGCTGGTGGAGAT 2340
Db 2281 CACGACACTCGGGTGGCAGTGGCTACCTACTTTGCTGAGCTCAAGCTGGTGGAGAT 2340
Qy 2341 GCTGTGACGCTGCTTCGCGACCACTGCTCGCGACCACTCCAGCCATCCCTCTCTGCC 2400
Db 2341 GCTGTGACGCTGCTTCGCGACCACTGCTCGCGACCACTCCAGCCATCCCTCTCTGCC 2400
Qy 2401 ATTGTGCTGATTCGCGCACACCGCTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCT 2460
Db 2401 ATTGTGCTGATTCGCGCACACCGCTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCT 2460
Qy 2461 GACCTCGAGCCGTAATGGAAGCTGTGGCGGAGCTGATACCTGCGCATTTGAGTCTGGAACC 2520
Db 2461 GACCTCGAGCCGTAATGGAAGCTGTGGCGGAGCTGATACCTGCGCATTTGAGTCTGGAACC 2520
Qy 2521 CCAGGCCAACCGGTCGCGTCTACCGCACCAAGATCCAGCGGACAGTGTCCCAACCTG 2580
Db 2521 CCAGGCCAACCGGTCGCGTCTACCGCACCAAGATCCAGCGGACAGTGTCCCAACCTG 2580
Qy 2581 CGTGACAGGCGCACCGCACTGGGCTTGCTGATCGCTTCGAGCTCATCGAAGACAACTAC 2640
Db 2581 CGTGACAGGCGCACCGCACTGGGCTTGCTGATCGCTTCGAGCTCATCGAAGACAACTAC 2640
Qy 2641 GCACCGGTTAATGATGCTGGGACGCGCCAAACCAAGGTACCCCATCTCCAGGTTGTT 2700
Db 2641 GCACCGGTTAATGATGCTGGGACGCGCCAAACCAAGGTACCCCATCTCCAGGTTGTT 2700
Qy 2701 GGGCACTCGCACTCCACTGGTGGTGGGCTGTAGATCCAGCAGACTTTGCTGCGAGAC 2760
Db 2701 GGGCACTCGCACTCCACTGGTGGTGGGCTGTAGATCCAGCAGACTTTGCTGCGAGAC 2760

Qy 2761 CCACAAAAGTACGACATCCAGACTCTGTGTATCGCGTTCCTGCGCGCGAGCTTGGTAAC 2820
Db 2761 CCACAAAAGTACGACATCCAGACTCTGTGTATCGCGTTCCTGCGCGCGAGCTTGGTAAC 2820
Qy 2821 CTTCCAGGTGGCTGGCGAGAACCACTGCGCACCGCGCACTGGAAGCCGCTCCGAAGGC 2880
Db 2821 CTTCCAGGTGGCTGGCGAGAACCACTGCGCACCGCGCACTGGAAGCCGCTCCGAAGGC 2880
Qy 2881 AAGGCACCTCTGACGGAAGTTCTGAGGAAGAGCAGCGCACCTCGACGCTGATGATTCC 2940
Db 2881 AAGGCACCTCTGACGGAAGTTCTGAGGAAGAGCAGCGCACCTCGACGCTGATGATTCC 2940
Qy 2941 AAGGAAGCTCGAAACAGCTTCAACCGCTCTGTTCCGGAAGCCAAACCGAAGAGTTCCTC 3000
Db 2941 AAGGAAGCTCGAAACAGCTTCAACCGCTCTGTTCCGGAAGCCAAACCGAAGAGTTCCTC 3000
Qy 3001 GAGCACCGTCCGCGCTTCGGCAACACCTCTGCGCTGATGATCGTGAATTTCTTACGGA 3060
Db 3001 GAGCACCGTCCGCGCTTCGGCAACACCTCTGCGCTGATGATCGTGAATTTCTTACGGA 3060
Qy 3061 CTGCTCGAGGCGCGGAGACTTTGATCCGCTCGCAGATGTGGCGACCCCACTGCTTGT 3120
Db 3061 CTGCTCGAGGCGCGGAGACTTTGATCCGCTCGCAGATGTGGCGACCCCACTGCTTGT 3120
Qy 3121 CGCTCGATGCGATCTCTGAGCCAGACGATAGGGTATGCGCAATGTTGTGGCCAAAGTC 3180
Db 3121 CGCTCGATGCGATCTCTGAGCCAGACGATAGGGTATGCGCAATGTTGTGGCCAAAGTC 3180
Qy 3181 AAGCGCCAGATCCGCCAATGCGTGTGCGACGCTCCGTTGAGTCTGTCACCGCAACC 3240
Db 3181 AAGCGCCAGATCCGCCAATGCGTGTGCGACGCTCCGTTGAGTCTGTCACCGCAACC 3240
Qy 3241 GCAGAAAAGGACAGATTCCTCCAAAGGGCCATGTTGCTGCACCATTCGCTGGTGTGTC 3300
Db 3241 GCAGAAAAGGACAGATTCCTCCAAAGGGCCATGTTGCTGCACCATTCGCTGGTGTGTC 3300
Qy 3301 ACTGTGACTGTTGCTGAAGGTGATGAGTCAAGGCTGGAGATCGAGTCGCAATCATCGAG 3360
Db 3301 ACTGTGACTGTTGCTGAAGGTGATGAGTCAAGGCTGGAGATCGAGTCGCAATCATCGAG 3360
Qy 3361 GCTATGAGATGGAAGCAACATCACTGTTTGAAGGCAAGATTGAACGGTGTG 3420
Db 3361 GCTATGAGATGGAAGCAACATCACTGTTTGAAGGCAAGATTGAACGGTGTG 3420
Qy 3421 GTTCTGCTCAACGAAGGTGGAAGGTGGGACTGATCGTCTGTTTCTCTAA 3474
Db 3421 GTTCTGCTCAACGAAGGTGGAAGGTGGGACTGATCGTCTGTTTCTCTAA 3474

RESULT 2

AAA47533

ID AAA47533 standard; DNA; 3621 BP.

XX AAA47533;

XX AC AC

XX DT 20-OCT-2000 (first entry)

XX XX

Sequence encoding pyruvate carboxylase of C. glutamicum.

DE DE

XX XX

Pyruvate carboxylase; expression; amino acid biosynthesis; lysine; glutamic acid; oxaloacetate; fermentation; biosynthesis; ds.

XX XX

OS Corynebacterium glutamicum.

XX Key

FH Location/Qualifiers

CDS 199

FT /tag= a

FT /product= Pyruvate carboxylase

FT /transl_except= (pos:199..202, aa:Met)

XX XX

PN W0200039305-A1.

XX XX

PD 06-JUL-2000.
 XX 23-DEC-1998; 98WO-US27301.
 XX 23-DEC-1998; 98WO-US27301.
 PR (SINS/) SINSKEY A J.
 PA (LESS/) LESSARD P A.
 PA (WILL/) WILLIS L B.
 XX Sinskey AJ, Lessard PA, Willis LB;
 PI WPI; 2000-465746/40.
 XX P-PSDB; AAB01436.
 DR Novel polynucleotides encoding Corynebacterium glutamicum pyruvate
 PT carboxylase useful for industrial fermentation processes comprises a
 PT specific nucleotide sequence
 XX Claim 3; Fig 1; 51pp; English.
 XX The pyruvate carboxylase of Corynebacterium glutamicum can be used
 CC for producing amino acids, preferably lysine and glutamic acid in
 CC industrial fermentations and for replenishing oxaloacetate consumed
 CC for biosyntheses during growth. By incorporating the pyruvate
 CC carboxylase gene in expression vectors levels of expression can be
 CC 2 - 20 fold higher than in Corynebacterium glutamicum.
 XX Sequence 3621 BP; 782 A; 1037 C; 1002 G; 800 T; 0 other;
 SQ
 Query Match 97.8%; Score 3398.8; DB 21; Length 3621;
 Best Local Similarity 98.6%; Pred. No. 0;
 Matches 3427; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
 Qy 1 GTCACTGCTATCAACCTTGGCGGTCTCTTTGTTGAAGGAATAATTAATCTAGTGTGCACT 60
 Db 148 GTCACTGCTATCAACCTTGGCGGTCTCTTTGTTGAAGGAATAATTAATCTAGTGTGCACT 207
 Qy 61 CACACATCTTCAACGGTTCAGCAATTCAAAAAGATCTTGTAGCAAAACCGCGGCGAAATC 120
 Db 208 CACACATCTTCAACGGTTCAGCAATTCAAAAAGATCTTGTAGCAAAACCGCGGCGAAATC 267
 Qy 121 GCGTCCGTCTTCCGTGCGACATTCGAACCGGTGCGAGCCAGGTAGCTATTATACCCC 180
 Db 268 GCGTCCGTCTTCCGTGCGACATTCGAACCGGTGCGAGCCAGGTAGCTATTATACCCC 327
 Qy 181 GGTGAAGATCGGGATCATTCACCGCTCTTTTGTCTTGAAGCTGTCGGCATTTGTACT 240
 Db 328 GTGAAGATCGGGATCATTCACCGCTCTTTTGTCTTGAAGCTGTCGGCATTTGTACT 387
 Qy 241 GAAGGCTCACAGTCAAGGCTACCTGGACATCGATGAATTAATCGGTGACGCTTAAAAA 300
 Db 388 GAAGGCTCACAGTCAAGGCTACCTGGACATCGATGAATTAATCGGTGACGCTTAAAAA 447
 Qy 301 GTTAAAGCAGATGCTATTTACCGGGATATGGCTTCTCTGTCTGAAATGCCCCAGCTTGCC 360
 Db 448 GTTAAAGCAGATGCTATTTACCGGGATACGGCTTCTCTGTCTGAAATGCCCCAGCTTGCC 507
 Qy 361 CGCGAGTGCAGGAAACGSCATTAATTTATTTGCCCCAAACCCAGAGGTTCTTGTATCTC 420
 Db 508 CGCGAGTGCAGGAAACGSCATTAATTTATTTGCCCCAAACCCAGAGGTTCTTGTATCTC 567
 Qy 421 ACCGGTATTAAGTCTCTGCGGTAAACCGCGGAAGAGCTGTCTCCAGATTTTGGCG 480
 Db 568 ACCGGTATTAAGTCTCTGCGGTAAACCGCGCGGAAGAGCTGTCTCCAGATTTTGGCG 627
 Qy 481 GAATCCACCCGAGCAAAACATCGATGATCGTGTAAAGCGCTGAAGCGCAGACTTAC 540
 Db 628 GAATCCACCCGAGCAAAACATCGATGATCGTGTAAAGCGCTGAAGCGCAGACTTAC 687
 Qy 541 CCATCTTTTGAAGCAGTTCCGCTGGGAGCGGATCGGCTTCTGTTCTTCA 600
 Db 688 CCCATCTTTGAGGCGAGTTCCCGGTGTGGCGGACCGGATGCGTTTGTGCTTCA 747

Qy 601 CCTGATGAGCTCGCAAAATGGCAACAGAAAGCATCTCGTGAAGCTGAAGCGCATTCGGC 660
 Db 748 CCTGATGAGCTTCGCAAAATAGCAACAGAAAGCATCTCGTGAAGCTGAAGCGCATTCGGC 807
 Qy 661 GACGTTCCGTATATGCGAAGCTGCTGATTAACCCCGACGACATTTGAAGTGCAGATC 720
 Db 808 GATGCGCGGTATATGCGAAGCTGCTGATTAACCCCGACGATTTGAAGTGCAGATC 867
 Qy 721 CTGCGGATCCGACCTGAGAAAGTTGTACACCTTTATGAAGCTGCTCTCACTGAGGCT 780
 Db 868 CTGCGGATCACACTGAGAAAGTTGTACACCTTTATGAAGCTGCTCTCACTGAGGCT 927
 Qy 781 CGTCAACAAAAGATTTGCAAAATTTGCGCCAGCACAGCATTTGGATCCAGAACTCGGTAT 840
 Db 928 CGTCAACAAAAGATTTGCAAAATTTGCGCCAGCACAGCATTTGGATCCAGAACTCGGTAT 987
 Qy 841 CGCATTTGTCCGATGCGATAAGTTCTGCGCTCTCATTTGGTTTACAGGGCGCGGAAACC 900
 Db 988 CGCATTTGTCCGATGCGATAAGTTCTGCGCTCTCATTTGGTTTACAGGGCGCGGAAACC 1047
 Qy 901 GTGGAATTTCTGTCGATGAAAAGGCAACACAGTTTTCATCGAAATGAACCCAGTATC 960
 Db 1048 GTGGAATTTCTGTCGATGAAAAGGCAACACAGTTTTCATCGAAATGAACCCAGTATC 1107
 Qy 961 CAGGTTGAGCACACCGTGTGAGAAAGTTCAGGAGTGGACCTGCTGAAGCGCGAGATG 1020
 Db 1108 CAGGTTGAGCACACCGTGTGAGAAAGTTCAGGAGTGGACCTGCTGAAGCGCGAGATG 1167
 Qy 1021 CGTTGGCTGCTGGTGCACCTTGAAGGAATTTGGTCTGACCAAGATGAAGTCAAGACC 1080
 Db 1168 CGTTGGCTGCTGGTGCACCTTGAAGGAATTTGGTCTGACCAAGATGAAGTCAAGACC 1227
 Qy 1081 CACGGTGCAGCATGCGAGTCCGCATCAACCGGAAGATTCAAAACAGCGCTTCGCGCCA 1140
 Db 1228 CACGGTGCAGCATGCGAGTCCGCATCAACCGGAAGATTCAAAACAGCGCTTCGCGCCA 1287
 Qy 1141 GATACCGGAATCAACCGCTACCGCTACCGGCGGAGCTGGCGTTCTGCTTCAACCGT 1200
 Db 1288 GATACCGGAATCAACCGCTACCGCTACCGGCGGAGCTGGCGTTCTGCTTCAACCGT 1347
 Qy 1201 GCAGCTCAGCTCGGTGCGCAAAATCAACCGCACACTTTGATCTCCATGCTGTGAAATGACC 1260
 Db 1348 GCAGCTCAGCTCGGTGCGCAAAATCAACCGCACACTTTGATCTCCATGCTGTGAAATGACC 1407
 Qy 1261 TGCGTGGTTCCGACTTTGAAACTGCTGTTGCTGCGTGCACAGCGCGTTGGCTGAGTTC 1320
 Db 1408 TGCGTGGTTCCGACTTTGAAACTGCTGTTGCTGCGTGCACAGCGCGTTGGCTGAGTTC 1467
 Qy 1321 ACCGTGCTGGTGTGCAACCAATGTTGCTGCGTGGTTCGCGGAGAGGAC 1380
 Db 1468 ACCGTGCTGGTGTGCAACCAATGTTGCTGCGTGGTTCGCGGAGAGGAC 1527
 Qy 1381 TTCACTTCCAGCGCATTCGCCACCGGATTTATCGGCGATCACCCAACCTCTTCAGGCT 1440
 Db 1528 TTCACTTCCAGCGCATTCGCCACCGGATTTATCGGCGATCACCCGACCTCTCTCAGGCT 1587
 Qy 1441 CCACCTCGGATGATGAGCAGGAGCATCTCTGATTTACTTGGCAGATGTCACTCGTGAAAC 1500
 Db 1588 CCACCTCGGATGATGAGCAGGAGCATCTCTGATTTACTTGGCAGATGTCACTCGTGAAAC 1647
 Qy 1501 AAGCTCATGCTGCTGCTCCAAAGGATTTGCGACCAACCAATCGATAAGCTGCCAACATC 1560
 Db 1648 AAGCTCATGCTGCTGCTCCAAAGGATTTGCGACCTCTTATCGATAAGCTGCCAACATC 1707
 Qy 1561 AAGATCTGCCACCTGCCACCGGTTCCCGTACCGCTTGAAGCAGCTTGGCCCGAGCGCG 1620
 Db 1708 AAGATCTGCCACCTGCCACCGGTTCCCGTACCGCTTGAAGCAGCTTGGCCCGAGCGCG 1767
 Qy 1621 TTTGCTCGTATCTCGTGAAGCAGGACGACTGGGAGTGTACTGATACCACTTCCGCGAT 1680
 Db 1768 TTTGCTCGTATCTCGTGAAGCAGGACGACTGGGAGTGTACTGATACCACTTCCGCGAT 1827

```
QY 1681 GCACACGACTCTTTGCTTGCAGCCGAGTCGCTCATTTGCACTGAAGCCCTGCGCAGAG 1740
DB 1828 GCACACGACTCTTTGCTTGCAGCCGAGTCGCTCATTTGCACTGAAGCCCTGCGCAGAG 1887
QY 1741 GCCGTGCGAAAGCTGACTCTTGAGCTTTTGTCCGTGAGAGGCTGCGGCGCGCACCTTAC 1800
DB 1888 GCCGTGCGAAAGCTGACTCTTGAGCTTTTGTCCGTGAGAGGCTGCGGCGCGCACCTTAC 1947
QY 1801 GATGTGGGATGCGTTTCTCTTTGAGATCCGTGGGACAGGCTCGAGAGCTGCGGAG 1860
DB 1948 GATGTGGGATGCGTTTCTCTTTGAGATCCGTGGGACAGGCTCGAGAGCTGCGGAG 2007
QY 1861 GCGATGCGGAATGTAACATTTCAGATGCTGCTCGCGCCGCAACACGTTGGGATACAC 1920
DB 2008 GCGATGCGGAATGTAACATTTCAGATGCTGCTCGCGCCGCAACACGTTGGGATACAC 2067
QY 1921 CCGTACCAGACTCCGCTGCGCGCGCTTTGTTAAGGAAGCTGCGAGCTCCGCGGTGGAC 1980
DB 2068 CCGTACCAGACTCCGCTGCGCGCGCTTTGTTAAGGAAGCTGCGAGCTCCGCGGTGGAC 2127
QY 1981 ATCTTCGCAATCTCGAGCGCTTAAAGAGTCTTCCAGATGCGTCCAGCAATCGACGCA 2040
DB 2128 ATCTTCGCAATCTCGAGCGCTTAAAGAGTCTTCCAGATGCGTCCAGCAATCGACGCA 2187
QY 2041 GTCTTGAGACCAACCGCGGTAGCGAGTGCTATGCGTATTTGTTGATCTCTCT 2100
DB 2188 GTCTTGAGACCAACCGCGGTAGCGAGTGCTATGCGTATTTGTTGATCTCTCT 2247
QY 2101 GATCCAAATGAAAGCTCTACACCTCGATTAACCTTAAAGATGGCAGAGGATCGTC 2160
DB 2248 GATCCAAATGAAAGCTCTACACCTCGATTAACCTTAAAGATGGCAGAGGATCGTC 2307
QY 2161 AAGTCTGGGCTCACTTTGGGCAATTAAGATATGCTGCTTGCCTCGCCAGCTGGC 2220
DB 2308 AAGTCTGGGCTCACTTTGGGCAATTAAGATATGCTGCTTGCCTCGCCAGCTGGC 2367
QY 2221 GTAACCAAGCTGTCACCGCACTGCGCGCTGAATTCGATCTGCTGCTGCTGCTGCTG 2280
DB 2368 GTAACCAAGCTGTCACCGCACTGCGCGCTGAATTCGATCTGCTGCTGCTGCTGCTGCA 2427
QY 2281 CACGACACTCGGGTGCGCACTTGGCTACCTTCTGCTGAGCTCAAGCTGGTGCAAT 2340
DB 2428 CACGACACTCGGGTGCGCACTTGGCTACCTTCTGCTGAGCTCAAGCTGGTGCAAT 2487
QY 2341 GCTGTTGAGGCTGCTTCGCAACCACTGCTGCGCACCACTCCAGCCATCCCTGTCTGCC 2400
DB 2488 GCTGTTGAGGCTGCTTCGCAACCACTGCTGCGCACCACTCCAGCCATCCCTGTCTGCC 2547
QY 2401 ATTGTTGCTGCATTCGCGCACACCGGTCGCGATACCGGTTTGAGCTCGAGGCTGTTCT 2460
DB 2548 ATTGTTGCTGCATTCGCGCACACCGGTCGCGATACCGGTTTGAGCTCGAGGCTGTTCT 2607
QY 2461 GACTTCGAGCCGCTACTGCGGAGCTGTGCGGAGCTGTACCTGCCATTTGAGTCTGGAACC 2520
DB 2608 GACTTCGAGCCGCTACTGCGGAGCTGTGCGGAGCTGTACCTGCCATTTGAGTCTGGAACC 2667
QY 2521 CAGGCCCAACCGGTGCGCTTACCGCAAGAAATCCAGCGCGAGCTGTGCTCCAACTG 2580
DB 2668 CAGGCCCAACCGGTGCGCTTACCGCAAGAAATCCAGCGCGAGCTGTGCTCCAACTG 2727
QY 2581 CGTGCAAGGCCCAACCGCACTGGGCTTGTGATCGCTTCGAGCTCATCGAAGCAACTAC 2640
DB 2728 CGTGCAAGGCCCAACCGCACTGGGCTTGTGATCGCTTCGAGCTCATCGAAGCAACTAC 2787
QY 2641 GCAGCCGTTAATGAGATGCTGGGACGCCCAAGAGTCAACCCATCCCTCCAAAGTTGTT 2700
DB 2788 GCAGCCGTTAATGAGATGCTGGGACGCCCAAGAGTCAACCCATCCCTCCAAAGTTGTT 2847
QY 2701 GGCAGCTCGCACTCCACCTGGTTGGTGGCGGTGTAGATCCAGCAGACTTTGCTGCAGAC 2760
DB 2848 GGCAGCTCGCACTCCACCTGGTTGGTGGCGGTGTAGATCCAGCAGACTTTGCTGCCGAT 2907
QY 2761 CCACAAAGTACGACATCCAGACTCTGTCTATCGCGTTCTGCGCGCGAGCTTGGTAAC 2820
```

```
DB 2308 CCACAAAGTACGACATCCAGACTCTGTCTATCGGTTCTCGCGCGAGCTTGGTAAC 2967
QY 2821 CTTCCAGGTGGCTGGCCAGAACCACTCGGCACCCCGCGCACTTGAAGGCCGCTCCGAAGGC 2880
DB 2968 CTTCCAGGTGGCTGGCCAGAGCCACCTCGGCACCCCGCGCACTTGAAGGCCGCTCCGAAGGC 3027
QY 2881 AAGCAGCTCTGACGGAAGTTCTTGAGAAAGAGAGAGGCGGCACTTCGAGCGTGTGATTC 2940
DB 3028 AAGCAGCTCTGACGGAAGTTCTTGAGAAAGAGAGAGGCGGCACTTCGAGCGTGTGATTC 3087
QY 2941 AAGGAACGTCGCAACAGCTCAACCGCTGCTGTTCCGAGAGCCAAACGGAAGAGTTCCTC 3000
DB 3088 AAGGAACGTCGCAATAGCTCAACCGCTGCTGTTCCGAGAGCCAAACGGAAGAGTTCCTC 3147
QY 3001 GAGCAGCTGCGCGCTTCGGCAACACCTCTGCGCTGATGATCGTGAATTCCTTCAAGGA 3060
DB 3148 GAGCAGCTGCGCGCTTCGGCAACACCTCTGCGCTGATGATCGTGAATTCCTTCAAGGA 3207
QY 3061 CTGCTGAGGCGCGCGAGACTTTGATCCGCTGCGCAGATGTGCGCACCCCACTGCTGTT 3120
DB 3208 CTGCTGAGGCGCGCGAGACTTTGATCCGCTGCGCAGATGTGCGCACCCCACTGCTGTT 3267
QY 3121 CGCTGAGTGGATCTCTGAGCGAGAGATGAAGGTTATGCGCAATGTTGTGGCCAAAGTC 3180
DB 3268 CGCTGAGTGGATCTCTGAGCGAGAGATGAAGGTTATGCGCAATGTTGTGGCCAAAGTC 3327
QY 3181 AAGCGCCAGATCCGCGCCAAATGCGTGTGCGTGAACCGCTTCTGTCACCGCAACC 3240
DB 3328 AAGCGCCAGATCCGCGCCAAATGCGTGTGCGTGAACCGCTTCTGTCACCGCAACC 3387
QY 3241 GCAGAAAGCAGATCTCTCCAAAGGGCCATGTTGCTGCACCAATTCGCTGCTGTTGTC 3300
DB 3388 GCAGAAAGCAGATCTCTCCAAAGGGCCATGTTGCTGCACCAATTCGCTGCTGTTGTC 3447
QY 3301 ACTGTGACTCTGCTGAAGGTGATGAGTCAAGCTGGAGATGCAAGTCCCAATCATCGAG 3360
DB 3448 ACCGTGACTCTGCTGAAGGTGATGAGTCAAGCTGGAGATGCAAGTCCCAATCATCGAG 3507
QY 3361 GCTATGAAGTGAAGCAACAATCACTGCTTCTGTTGACCGCAAGATTGAACGGTGTG 3420
DB 3508 GCTATGAAGTGAAGCAACAATCACTGCTTCTGTTGACCGCAAAATCGATCGGTTGTG 3567
QY 3421 GTTCTGCTGCAACGAAAGTGAAGTGGCAGCTTGATGCTGCTGTTTCTTAA 3474
DB 3568 GTTCTGCTGCAACGAAAGTGAAGTGGCAGCTTGATGCTGCTGTTTCTTAA 3621
```

RESULT 3

AAF32165

ID AAF32165 standard; DNA; 3621 BP.

XX AAF32165;

XX AC

XX XX

DT 12-APR-2001 (first entry)

XX XX

DE Corynebacterium glutamicum pyruvate carboxylase coding sequence.

XX Pyruvate carboxylase; anaplerotic pathway; industrial fermentation;

XX oxaloacetate; ds.

XX OS

XX Corynebacterium glutamicum.

XX PN

XX US6171833-B1.

XX XX

PD 09-JAN-2001.

XX XX

PF 23-DEC-1998; 98US-0220081.

XX XX

PR 23-DEC-1998; 98US-0220081.

XX XX

PA (MASI) MASSACHUSETTS INST TECHNOLOGY.

XX XX

PI Sinskey AJ, Lessard PA, Willis LB;
XX WPI: 2001-122330/13.
DR P-PSDB; AAB67129.
XX Novel nucleic acid encoding pyruvate carboxylase from *Corynebacterium*
PT glutamicum, for replenishing oxaloacetate consumed during lysine and
PT glutamic acid production in industrial fermentations -
XX Claim 2; Column 23-30; 29pp; English.
PS The present invention provides the protein and coding sequences of the
XX *Corynebacterium glutamicum* pyruvate carboxylase protein. This is an
CC enzyme in the anaplerotic pathway. It can be used in the replenishment of
CC oxaloacetate consumed during lysine and glutamic acid production in
CC industrial fermentation.
XX Sequence 3621 BP; 782 A; 1037 C; 1002 G; 800 T; 0 other;
SQ Query Match 97.8%; Score 3398.8; DB 22; Length 3621;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3427; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 1 GTGACTGCTATCACGCTTGGCGGCTCTTGTGTAAGGAATAATTACTCTAGTGTGACT 60
DB 148 GTGACTGCTATCACGCTTGGCGGCTCTTGTGTAAGGAATAATTACTCTAGTGTGACT 207
QY 61 CACACATCTTCAACGGCTTCACGATTCAAAAGATCTTGGTAGCAAAACCGCGGGAATC 120
DB 208 CACACATCTTCAACGGCTTCACGATTCAAAAGATCTTGGTAGCAAAACCGCGGGAATC 267
QY 121 GCGGTCGCTGTTCCGTCGACGACATCGAAACCGGTCGAGCCAGGTAGCTATTTACCCC 180
DB 268 GCGGTCGCTGTTCCGTCGACGACATCGAAACCGGTCGAGCCAGGTAGCTATTTACCCC 327
QY 181 CGTGAAGATCGGGGATCAATCCACCGCTCTTTTGTCTGAAGCTGTCGCGATGTGACT 240
DB 328 CGTGAAGATCGGGGATCAATCCACCGCTCTTTTGTCTGAAGCTGTCGCGATGTGACT 387
QY 241 GAAGGCTCACAGTCAAGCGTACCTGGACATCGAATGAAATTTATCGGTGAGCTAAAAA 300
DB 388 GAAGGCTCACAGTCAAGCGTACCTGGACATCGAATGAAATTTATCGGTGAGCTAAAAA 447
QY 301 GTTAAAGCAGATGCTATTTACCGGGATATGGCTTCTGTCGTAAGATGCCAGCTTGCC 360
DB 448 GTTAAAGCAGATGCTATTTACCGGGATATGGCTTCTGTCGTAAGATGCCAGCTTGCC 507
QY 361 GCGAGTGGCGGAAACCGGATTAATTTATGGCCCAACCCAGAGGTTCTTGATCTC 420
DB 508 GCGAGTGGCGGAAACCGGATTAATTTATGGCCCAACCCAGAGGTTCTTGATCTC 567
QY 421 ACCGTTGATAGTCTCGTCGGTAAACCGCGGAAAGAGGCTGCTGCCAGGTTTGGCG 480
DB 568 ACCGTTGATAGTCTCGTCGGTAAACCGCGGAAAGAGGCTGCTGCCAGGTTTGGCG 627
QY 481 GAATPCCACCCGAGCAAAACATCGATGACATCGTTAAAGCGCTGAAGCCAGACTTAC 540
DB 628 GAATPCCACCCGAGCAAAACATCGATGATGATCGTTAAAGCGCTGAAGCCAGACTTAC 687
QY 541 CCCATCTTTGTAAGGACGTTGCGGTTGTCGCGACCGGATATGCGCTTTGTTCTTCA 600
DB 688 CCCATCTTTGTAAGGACGTTGCGGTTGTCGCGACCGGATATGCGCTTTGTTCTTCA 747
QY 601 CCTGATGAGTCTCGCAAAATGGCAACAGAGCATCTCTGTAAGCTGAAGCGGCAATTCGGC 660
DB 748 CCTGATGAGTCTCGCAAAATAGCAACAGAGCATCTCTGTAAGCTGAAGCGGCTTTCGGC 807
QY 661 GACGGTTTCGTATATCGAACTGCTGATTAACCCCGCAGCATTTGAAGTCAGATC 720
DB 808 GATGGCGGATATATGTCGAACTGCTGATTAACCCCGCAGCATTTGAAGTCAGATC 867
QY 721 CTTGGGATCGCACTCGAGAGTGTGACACCTTTATGACCTGTGCTCACTGACGGT 780

DB 868 CTTGGCGATCACACTGGAGAGTTGTACACCTTTATGAACTGCTGCTCACTGACGGT 927
QY 781 CGTCAACCAAAAGTTGTGAAATTCGCCAGCACAGCATTTTGGATCCAGAACTCGCTGAT 840
DB 928 CGTCAACCAAAAGTTGTGAAATTCGCCAGCACAGCATTTTGGATCCAGAACTCGCTGAT 987
QY 841 CGATTTGTGCGGATGACGTAAGTTCTGCGCTCCATTTGGTTACCGAGGCGCGGGAACC 900
DB 988 CGATTTGTGCGGATGACGTAAGTTCTGCGCTCCATTTGGTTACCGAGGCGCGGGAACC 1047
QY 901 GTGGAATCTTGTGATGAAAGGGCAACACGTTTTCATCGAAATGAACCCACGATC 960
DB 1048 GTGGAATCTTGTGATGAAAGGGCAACACGTTTTCATCGAAATGAACCCACGATC 1107
QY 961 CAGGTTGAGCACACCGTGTACTGAAGAAGTCAACGAGGTGGACCTGTGTGAAGGCGCAGATG 1020
DB 1108 CAGGTTGAGCACACCGTGTACTGAAGAAGTCAACGAGGTGGACCTGTGTGAAGGCGCAGATG 1167
QY 1021 CGTTTGGCTGCTGGTGCAACCTTTGAAGGAATTTGGGTCTGACCCCAAGATAAGATCAAGACC 1080
DB 1168 CGTTTGGCTGCTGGTGCAACCTTTGAAGGAATTTGGGTCTGACCCCAAGATAAGATCAAGACC 1227
QY 1081 CACGTTGACGACTGACGTCGCGCATCACCGGAAGATCAAAACACCGGCTTCGGCCCA 1140
DB 1228 CACGTTGACGACTGACGTCGCGCATCACCGGAAGATCAAAACACCGGCTTCGGCCCA 1287
QY 1141 GATACCGGACTATCACCGGTACCGCTCACCGCGGAGCTGGCGTTTCGTTGACGGT 1200
DB 1288 GATACCGGACTATCACCGGTACCGCTCACCGCGGAGCTGGCGTTTCGTTGACGGT 1347
QY 1201 GCGACTCAGCTCGGTGGGAAATCAACCGGACACTTTGATCTCATGCTGGTGAATGACC 1260
DB 1348 GCGACTCAGCTCGGTGGGAAATCAACCGGACACTTTGATCTCATGCTGGTGAATGACC 1407
QY 1261 TGCCTGGTTCCGACTTTGAAATCTGTTGCTGTGTGTCACAGCGCGGCTTGGCTGAGTTC 1320
DB 1408 TGCCTGGTTCCGACTTTGAAATCTGTTGCTGTGTGTCACAGCGCGGCTTGGCTGAGTTC 1467
QY 1321 ACCGTGTCTGTTGTTCAACCAACATTTGTTTCTTGGCTGCTGTCGCGGAAGAGGAC 1380
DB 1468 ACCGTGTCTGTTGTTCAACCAACATTTGTTTCTTGGCTGCTGTCGCGGAAGAGGAC 1527
QY 1381 TTCACTTCCAAAGCGCATCGCCACCGGATTTATCGGCGATCACCCACACCTCTCTTCAGGCT 1440
DB 1528 TTCACTTCCAAAGCGCATCGCCACCGGATTTATTTGCCGATCACCCGACCTCTCTTCAGGCT 1587
QY 1441 CCACCTGCGGATGATGACGAGGAGCGCATCTCTGGATTTACTTGGCAGATGTCAACGTTGAAC 1500
DB 1588 CCACCTGCTGATGATGACGAGGAGCGCATCTCTGGATTTACTTGGCAGATGTCAACGTTGAAC 1647
QY 1501 AAGCCTCATGTTGTTGCTCCAAAGGATGTTGAGCACCAATCGATAGCTGCCAAATC 1560
DB 1648 AAGCCTCATGTTGTTGCTCCAAAGGATGTTGAGCTCTTATCGATAGCTGCCAAATC 1707
QY 1561 AAGATCTGCACTGCGACCGGTTCCGTTGACCGCTGACAGGCTTGGCGCAGCGCG 1620
DB 1708 AAGATCTGCACTGCGACCGGTTCCGTTGACCGCTGACAGGCTTGGCGCAGCGCG 1767
QY 1621 TTTGCTCGTATCTCGTGAGCAGGACGCACTGGCAGTTACTGATPACACCTTTCGCGAT 1680
DB 1768 TTTGCTCGTATCTCGTGAGCAGGACGCACTGGCAGTTACTGATPACACCTTTCGCGAT 1827
QY 1681 GCACACAGTCTTTTGTGTCGACCCGAGTCCGCTCATTTGCACTGAAGCCTCGCGCAGAG 1740
DB 1828 GCACACAGTCTTTTGTGTCGACCCGAGTCCGCTCATTTGCACTGAAGCCTCGCGCAGAG 1887
QY 1741 GCCGTGCAAAAGCTGACTCTCTGAGCTTTTGTCTGAGGCTTGGGGCGGCGGACCTTAC 1800
DB 1888 GCCGTGCAAAAGCTGACTCTCTGAGCTTTTGTCTGAGGCTTGGGGCGGCGGACCTTAC 1947
QY 1801 GATGTGCGATGCGTTTCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGGAG 1860
DB 1948 GATGTGCGATGCGTTTCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGGAG 2007

Qy	1861	GC	GATG	CCG	GAATCT	TAACA	CAATT	CAG	ATG	CTG	CTCTT	CG	CGG	CCG	CAAC	AC	CGT	GGG	ATA	CACC	1920
Db	2008	GC	GATG	CCG	GAATCT	TAACA	CAATT	CAG	ATG	CTG	CTCTT	CG	CGG	CCG	CAAC	AC	CGT	GGG	ATA	CACC	2067
Qy	1921	CC	GT	TACC	CG	AGACT	CCG	CTCT	CG	CGG	CGT	TTG	T	T	T	T	T	T	T	T	1980
Db	2068	CC	GT	TACC	CG	AGACT	CCG	CTCT	CG	CGG	CGT	TTG	T	T	T	T	T	T	T	T	2127
Qy	1981	AT	CTT	CCG	CACT	TTT	CG	AC	CGG	CTT	AA	CG	AGT	CT	CCC	AG	ATG	CGT	CG	CA	2040
Db	2128	AT	CTT	CCG	CACT	TTT	CG	AC	CGG	CTT	AA	CG	AGT	CT	CCC	AG	ATG	CGT	CG	CA	2187
Qy	2041	GT	CCT	TG	GAG	AC	CA	AAC	CGG	GGT	TAG	CC	GAG	TGG	CT	TAT	TGG	CTT	AT	TT	2100
Db	2188	GT	CCT	TG	GAG	AC	CA	AAC	CGG	GGT	TAG	CC	GAG	TGG	CT	TAT	TGG	CTT	AT	TT	2247
Qy	2101	GAT	CC	AAAT	G	AA	AGCT	CTT	AC	AC	CC	TGG	AT	T	ACT	T	AA	AG	TGG	CA	2160
Db	2248	GAT	CC	AAAT	G	AA	AGCT	CTT	AC	AC	CC	TGG	AT	T	ACT	T	AA	AG	TGG	CA	2307
Qy	2161	AA	GT	CTGG	CG	CT	C	A	C	A	TT	CTGG	CC	AT	T	AA	G	A	T	T	2220
Db	2308	AA	GT	CTGG	CG	CT	C	A	C	A	TT	CTGG	CC	AT	T	AA	G	A	T	T	2367
Qy	2221	GT	AA	CC	AA	GCT	TGGT	CA	CCG	CA	CTG	CG	CG	CGT	GA	AT	T	CG	AT	T	2280
Db	2368	GT	AA	CC	AA	GCT	TGGT	CA	CCG	CA	CTG	CG	CG	CGT	GA	AT	T	CG	AT	T	2427
Qy	2281	C	A	G	C	A	C	A	CTG	CGG	TGG	CC	AGT	T	G	CT	A	C	T	A	2340
Db	2428	C	A	G	C	A	C	A	CTG	CGG	TGG	CC	AGT	T	G	CT	A	C	T	A	2487
Qy	2341	G	CT	TT	T	G	A	CG	GT	CTT	CC	G	A	C	A	CTG	CT	G	C	A	2400
Db	2488	G	CT	TT	T	G	A	CG	GT	CTT	CC	G	A	C	A	CTG	CT	G	C	A	2547
Qy	2401	A	T	T	G	T	G	CT	G	A	T	TCG	G	C	A	C	CCG	T	T	G	2460
Db	2548	A	T	T	G	T	G	CT	G	A	T	TCG	G	C	A	C	CCG	T	T	G	2607
Qy	2461	G	A	C	CT	CG	AG	CCG	T	A	CT	TGG	A	G	CTG	TG	CG	G	A	C	2520
Db	2608	G	A	C	CT	CG	AG	CCG	T	A	CT	TGG	A	G	CTG	TG	CG	G	A	C	2667
Qy	2521	C	A	G	G	C	C	A	A	C	G	G	T	C	G	T	A	C	C	G	2580
Db	2668	C	A	G	G	C	C	A	A	C	G	G	T	C	G	T	A	C	C	G	2727
Qy	2581	C	GT	G	C	A	G	G	C	C	A	C	G	G	T	C	G	T	A	C	2640
Db	2728	C	GT	G	C	A	G	G	C	C	A	C	G	G	T	C	G	T	A	C	2787
Qy	2641	G	C	A	G	CCG	TT	AA	T	G	A	T	G	CT	TGG	A	G	CC	CA	A	2700
Db	2788	G	C	A	G	CCG	TT	AA	T	G	A	T	G	CT	TGG	A	G	CC	CA	A	2847
Qy	2701	G	G	C	G	A	C	T	C	A	C	T	G	T	T	G	T	G	T	A	2760
Db	2848	G	G	C	G	A	C	T	C	A	C	T	G	T	T	G	T	G	T	A	2907

Qy	2941	AAGGAA	CGT	CGCAAC	GCCT	CTAA	CGCC	TGCT	GTGTTCC	CGAAG	CCAA	CGGAAG	AGTT	CTCT	3000
Db	3088	AAGGAA	CGT	CGCAAT	AGCT	CTAA	CGCC	TGCT	GTGTTCC	CGAAG	CCAA	CGGAAG	AGTT	CTCT	3147
Qy	3001	GAGCAC	CGT	CGCGCT	TCGG	CAAC	CT	CTCG	CTGG	ATGAT	CGTGA	ATTTCT	TTAC	CGGA	3060
Db	3148	GAGCAC	CGT	CGCGCT	TCGG	CAAC	CT	CTCG	CTGG	ATGAT	CGTGA	ATTTCT	TTAC	CGGC	3207
Qy	3061	CTGTG	CAGG	CGCG	CAGAC	CTTTG	ATC	CGCT	CGCC	AGAT	TGCG	CA	CCCA	CTGCT	3120
Db	3208	CTGTG	CAGG	CGCG	CAGAC	CTTTG	ATC	CGCT	CGCC	AGAT	TGCG	CA	CCCA	CTGCT	3267
Qy	3121	CGCTG	GAT	CGCAT	CTCT	GAGC	CAG	ACG	ATAAG	GGTAT	TGCG	CA	ATTTG	TGGC	3180
Db	3268	CGCTG	GAT	CGCAT	CTCT	GAGC	CAG	ACG	ATAAG	GGTAT	TGCG	CA	ATTTG	TGGC	3327
Qy	3181	AACGGC	CAGAT	TCG	CCCAAT	GTG	GTG	CGTGA	CCGCT	CGT	TGAGT	CTGT	CAC	CGCA	3240
Db	3328	AACGGC	CAGAT	TCG	CCCAAT	GTG	GTG	CGTGA	CCGCT	CGT	TGAGT	CTGT	CAC	CGCA	3387
Qy	3241	GCAGAA	AAGC	AGAT	TTC	TCC	CAAC	AGG	CGCAT	TTG	CTGC	CA	CCATTC	CGCT	3300
Db	3388	GCAGAA	AAGC	AGAT	TTC	TCC	CAAC	AGG	CGCAT	TTG	CTGC	CA	CCATTC	CGCT	3447
Qy	3301	ACTGTG	ACT	GTG	CTGA	AGTG	AT	GAGT	CAAG	GGCT	TGG	AGAT	TGC	AGT	3360
Db	3448	ACCGTGA	CTGTT	GCTGA	AGGT	GN	TG	AGGT	CA	AGG	CTT	GA	AGAT	TGC	3507
Qy	3361	GCTAT	GAAG	TGGA	AGCA	CAAT	CA	CTG	TCTT	CTGT	TGAC	GGCA	AGATT	GAA	3420
Db	3508	GCTAT	GAAG	TGGA	AGCA	CAAT	CA	CTG	TCTT	CTGT	TGAC	GGCA	AGATT	GAA	3567
Qy	3421	GTTCT	CTG	CTGAA	CGA	AGT	TGGA	AGT	GGCA	CTG	TAT	GAT	CTG	CTGTT	3474
Db	3568	GTTCT	CTG	CTGAA	CGA	AGT	TGGA	AGT	GGCA	CTG	TAT	GAT	CTG	CTGTT	3621

RESULT 4

RESUL 4
AAH68526

AAH68526
ID AAH68526 standard: DNA; 349980 BP.

XX
AC
AAH68526:

XX
DT 26-SEP-2001 (first entry)

XX
DE C glutamicum coding sequence fragment SEQ ID NO: 7061

XX
XX

KW Coryneform bacterium; amino
KW organic acid synthesis; de

KW organic acid synthesis; ds.
 YY

OS *Corynebacterium glutamicum*.
yy

PN EP1108790-A2.

PD 20-JUN-2001.

PF 18-DEC-2000; 2000EP-0127688.

PR 16-DEC-1999; 99JP-0377484.

PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.

XX PA (KYOW) KYOWA HAKKO KOGYO KK.

XX	Nakagawa S	Mizoguchi H	Ando S	Hayashi M	Ochiai Y	Yokoi H.
PT						

PI Nakagawa S, Mizoguchi H, Ando S, Hayas
PI Tateishi N, Senoh A, Ikeda M, Ozaki A.

XX
DP WBT. 2001 376031/40

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
Pr mutation point of a gene, measuring expression of a gene, analysing
Pr expression profile or pattern of a gene and identifying homologous gene

XX PS Disclosure; SEQ ID NO: 7061; 246pp + Sequence Listing; English.

XX CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium *Corynebacterium glutamicum*. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of *Corynebacterium glutamicum*, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from *Corynebacterium glutamicum*, and identifying a homologue of a gene derived
CC from *Corynebacterium glutamicum*. *Corynebacterium* bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.

XX SQ Sequence 349980 BP; 79274 A; 90638 C; 98727 G; 81341 T; 0 other;

Query Match 97.8%; Score 3398.8; DB 22; Length 349980;
Best Local Similarity 98.6%; Pred No. 0;
Matches 3427; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 1 GTGACTGCTATCACCTTGGCGGCTCTCTGTGGAAGGAATAATCTCTAGTGTGACT 60

Db 105160 GTGACTGCTATCACCTTGGCGGCTCTCTGTGGAAGGAATAATCTCTAGTGTGACT 105219

QY 61 CACACATCTTCAACGGTTCACAGATTCAAAAGATCTTGGTAGCAACCGCGGGAATC 120

Db 105220 CACACATCTTCAACGGTTCACAGATTCAAAAGATCTTGGTAGCAACCGCGGGAATC 105279

QY 121 GCGGTCCGCTCTTCCGTGACGACTCGAAACCGGTGAGCCAGGTAGCTATTATACCCC 180

Db 105280 GCGGTCCGCTCTTCCGTGACGACTCGAAACCGGTGAGCCAGGTAGCTATTATACCCC 105339

QY 181 CGTGAAGATCGGGGATCATTCACCGCTCTTTGCTTCTGAAGCTGTCGCGATTGCTACT 240

Db 105340 CGTGAAGATCGGGGATCATTCACCGCTCTTTGCTTCTGAAGCTGTCGCGATTGCTACT 105399

QY 241 GAAGGCTCACAGTCAAGCGGTACCTGGACATCGATGAAATTTATCGGTGAGCTTAAAAA 300

Db 105400 GAAGGCTCACAGTCAAGCGGTACCTGGACATCGATGAAATTTATCGGTGAGCTTAAAAA 105459

QY 301 GTTAAAGCAGATCTATTATACCGGGATATGGCTTCTGTCTGAAATGCCAGCTTGCC 360

Db 105460 GTTAAAGCAGATCTATTATACCGGGATATGGCTTCTGTCTGAAATGCCAGCTTGCC 105519

QY 361 CGCGATGCGCGAAGAACGCGATTACTTTTATTTGGCCCAACCCAGAGGTTCTTGATCTC 420

Db 105520 CGCGATGCGCGAAGAACGCGATTACTTTTATTTGGCCCAACCCAGAGGTTCTTGATCTC 105579

QY 421 ACCGGTGATAGCTCTCGTGGGTAAACCGCGAAGAGGCTGCTGCGAGTTTGGCG 480

Db 105580 ACCGGTGATAGCTCTCGTGGGTAAACCGCGAAGAGGCTGCTGCGAGTTTGGCG 105639

QY 481 GAATCCACCGAGCAAAACATCGATGATCGTTAAAGCGCTGAAGCCAGACTTAC 540

Db 105640 GAATCCACCGAGCAAAACATCGATGATCGTTAAAGCGCTGAAGCCAGACTTAC 105699

QY 541 CCCATCTTTGAAGCAGTTGCGGTGTGCGGACCGGTATGCGCTTTGTTTCTTCA 600

Db 105700 CCCATCTTTGAAGCAGTTGCGGTGTGCGGACCGGTATGCGCTTTGTTTCTTCA 105759

QY 601 CCGTAGAGCTCCGCAAAATGGCAACAGAGCATCTCGTGAAGCTGAAGCGGCAATCGGC 660

Db 105760 CCGTAGAGCTCCGCAAAATGGCAACAGAGCATCTCGTGAAGCTGAAGCGGCAATCGGC 105819

QY 661 GACGGTTCCGTATATGTGAAACGTGTGATTAAACCCCGACGACATTTGAAGTGCAGATC 720

Db 105820 GATGGCGCGGTATATGTGAAACGTGTGATTAAACCCCGACGACATTTGAAGTGCAGATC 105879

QY 721 CTTGGCGCATCGCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTCAGCGT 780

Db 105960 GATGGCGCATCGCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTCAGCGT 107019

Db 105880 CTTGGCGCATCGCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTCAGCGT 105939

QY 781 CGTCACCAAAAGTTGTCGAAATTCGGCAGCAGCAAGCAATTTGGATCCAGAACTCGGTGAT 840

Db 105940 CGTCACCAAAAGTTGTCGAAATTCGGCAGCAGCAAGCAATTTGGATCCAGAACTCGGTGAT 105999

QY 841 CGCATTTGTGCGGATGACGATTAAGTTCTGCGCTCCATTTGTTTACAGGGCGCGGGAACC 900

Db 106000 CGCATTTGTGCGGATGACGATTAAGTTCTGCGCTCCATTTGTTTACAGGGCGCGGGAACC 106059

QY 901 GTGGAATTTCTTGTGTCGATGAAAGGGCAACCAACGCTTTTTCATCGAAATGAACCCACGTATC 960

Db 106060 GTGGAATTTCTTGTGTCGATGAAAGGGCAACCAACGCTTTTTCATCGAAATGAACCCACGTATC 106119

QY 961 CAGGTTGAGCACACCTGTGACTGAGAAGTTCACCGAGGTGGACCTGTTGTAAGCGCGCAGATG 1020

Db 106120 CAGGTTGAGCACACCTGTGACTGAGAAGTTCACCGAGGTGGACCTGTTGTAAGCGCGCAGATG 106179

QY 1021 CGCTTGGCTGCTGCTGCAACCTTTGAAGGAATTTGGGTCTTGACCCCAAGATAAAGATCAAGACC 1080

Db 106180 CGCTTGGCTGCTGCTGCAACCTTTGAAGGAATTTGGGTCTTGACCCCAAGATAAAGATCAAGACC 106239

QY 1081 CACGGTGCAGCACTGACGTCGCGCATCAACCGGAAGATCCAAACMACGGCTTCGCGCCA 1140

Db 106240 CACGGTGCAGCACTGACGTCGCGCATCAACCGGAAGATCCAAACMACGGCTTCGCGCCA 106299

QY 1141 GATACCGGAATATGACCGGTACCGCTCACCGGGGAGCTGCGGTTTCGTTGACGCT 1200

Db 106300 GATACCGGAATATGACCGGTACCGCTCACCGGGGAGCTGCGGTTTCGTTGACGCT 106359

QY 1201 GCAGCTCAGCTCGGTGCGGAAATCAACCGCAGCACTTTGATCTCATGCTGCTGGAATAATGACC 1260

Db 106360 GCAGCTCAGCTCGGTGCGGAAATCAACCGCAGCACTTTGATCTCATGCTGCTGGAATAATGACC 106419

QY 1261 TGGCGTGTTCGCACTTTGAAACTGCTGTGTGCTGCAACGCGCGCTTGGCTGAGTTC 1320

Db 106420 TGGCGTGTTCGCACTTTGAAACTGCTGTGTGCTGCAACGCGCGCTTGGCTGAGTTC 106479

QY 1321 ACCGTGTCTGGTTTCAACCAACCATTTGGTTTCTTTCGCTGCTGCTGCGGGAAGAGAC 1380

Db 106480 ACCGTGTCTGGTTTCAACCAACCATTTGGTTTCTTTCGCTGCTGCTGCGGGAAGAGAC 106539

QY 1381 TTCACTTTCAAAGCGCATCGCCACCGGATTTATCGGCGATCACCCACACCTCTCTTCAGGCT 1440

Db 106540 TTCACTTTCAAAGCGCATCGCCACCGGATTTATCGGCGATCACCCACACCTCTCTTCAGGCT 106599

QY 1441 CCACCTGCGGATGATGAGCAGGAGCGCATCTCTGGATTTACTTTGGCAGATGTCACCGTGAAC 1500

Db 106600 CCACCTGCGGATGATGAGCAGGAGCGCATCTCTGGATTTACTTTGGCAGATGTCACCGTGAAC 106659

QY 1501 AAGCCTCATGCTGCTGCGTCCAAAGGATGTTGAGCAGCAACAAATCGATAAGCTGCCAAATC 1560

Db 106660 AAGCCTCATGCTGCTGCGTCCAAAGGATGTTGAGCAGCTCCTATCGATAAGCTGCCAAATC 106719

QY 1561 AAGGATCTGCACTGCGCAGCGGTTCCGCTGACCGCTGAGAGCAGCTTGGCCAGCGCG 1620

Db 106720 AAGGATCTGCACTGCGCAGCGGTTCCGCTGACCGCTGAGAGCAGCTTGGCCAGCGCG 106779

QY 1621 TTTGCTCGTATCTCGTGAAGAGCAGCACTGCGAGTTACTGATACCACTTCCGCGAT 1680

Db 106780 TTTGCTCGTATCTCGTGAAGAGCAGCACTGCGAGTTACTGATACCACTTCCGCGAT 106839

QY 1681 GCACACAGCTTTTGTCTGCGACCGCGAGTCCCTCATTCGCACTGAAGCCTGCGGCGAG 1740

Db 106840 GCACACAGCTTTTGTCTGCGACCGCGAGTCCCTCATTCGCACTGAAGCCTGCGGCGAG 106899

QY 1741 GCGGTCGCAAGCTGACTCTCTGAGCTTTTGTCCGTGGAGGCTTGGGGCGGCGGACCTTAC 1800

Db 106900 GCGGTCGCAAGCTGACTCTCTGAGCTTTTGTCCGTGGAGGCTTGGGGCGGCGGACCTTAC 106959

QY 1801 GATGTGCGATGCGGTTTCTCTTTTGAAGGATCCGTGGGACAGGCTCGACGAGCTCGCGAG 1860

Db 106960 GATGTGCGATGCGGTTTCTCTTTTGAAGGATCCGTGGGACAGGCTCGACGAGCTCGCGAG 107019

QY 1861 CGGATGCCGAATGTAAACATTCAGATGCTGTTCGGCGCCGCAACACCGTGGGATACACC 1920
|
|
|
Db 107020 CGGATGCCGAATGTAAACATTCAGATGCTGTTCGGCGCCGCAACACCGTGGGATACACC 107079
|
|
|
QY 1921 CCGTACCCAGACTCCGCTCTGCGCGCTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC 1980
|
|
|
Db 107080 CCGTACCCAGACTCCGCTCTGCGCGCTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC 107139
|
|
|
QY 1981 ATCTTCGCACTTTTCGACGCGCTTAAACGACGCTCTCCAGATGCGTCCAGCAATCGACGCA 2040
|
|
|
Db 107140 ATCTTCGCACTTTTCGACGCGCTTAAACGACGCTCTCCAGATGCGTCCAGCAATCGACGCA 107199
|
|
|
QY 2041 GTCTTCGAGACCAACACCGCGGTAGCCAGGTGGCTATGGCTTATTCGTGTGATCTCTCT 2100
|
|
|
Db 107200 GTCTTCGAGACCAACACCGCGGTAGCCAGGTGGCTATGGCTTATTCGTGTGATCTCTCT 107259
|
|
|
QY 2101 GATCCAAATGAAAGCTCTACACCTGGATTACTACCTTAAAGATGGCAGAGGAGATCGTC 2160
|
|
|
Db 107260 GATCCAAATGAAAGCTCTACACCTGGATTACTACCTTAAAGATGGCAGAGGAGATCGTC 107319
|
|
|
QY 2161 AAGTCTGCGCTCATTCTCGGCCATTAAAGGATATGGCTGTCTGCTTCGCCAGCTGCG 2220
|
|
|
Db 107320 AAGTCTGCGCTCATTCTCGGCCATTAAAGGATATGGCTGTCTGCTTCGCCAGCTGCG 107379
|
|
|
QY 2221 GTAACCAAGCTGTACCGCACTGCGCGTGGAAATTCGATCTGCCAGTGCACTGCAACACC 2280
|
|
|
Db 107380 GTAACCAAGCTGTACCGCACTGCGCGTGGAAATTCGATCTGCCAGTGCACTGCAACACC 107439
|
|
|
QY 2281 CACGACACTGCGGGTGGCCAGCTGCGTACCTACTTTGCTGCAGCTCAAGCTGGTGCAGAT 2340
|
|
|
Db 107440 CACGACACTGCGGGTGGCCAGCTGCGTACCTACTTTGCTGCAGCTCAAGCTGGTGCAGAT 107499
|
|
|
QY 2341 GCTGTGACGCTGTCTCGCAACCACTGTCTGGCACCACTCCAGGCAATCCCTGTCTGCC 2400
|
|
|
Db 107500 GCTGTGACGCTGTCTCGCAACCACTGTCTGGCACCACTCCAGGCAATCCCTGTCTGCC 107559
|
|
|
QY 2401 ATTGTGCTGATTCGGCGCAACCCGCTGCGGATACCCGGTTTGGAGCTCGAGGCTGTTTCT 2460
|
|
|
Db 107560 ATTGTGCTGATTCGGCGCAACCCGCTGCGGATACCCGGTTTGGAGCTCGAGGCTGTTTCT 107619
|
|
|
QY 2461 GACCTCGAGCCGTACTGGGAAGCTGTGCGGAGCTGTACCTGCCATTTGAGTCTGGAACC 2520
|
|
|
Db 107620 GACCTCGAGCCGTACTGGGAAGCTGTGCGGAGCTGTACCTGCCATTTGAGTCTGGAACC 107679
|
|
|
QY 2521 CCAGGCCCAACCGGTGCGCTTACCGCCACACGAAATCCAGCGGACACTTCTCCAACTG 2580
|
|
|
Db 107680 CCAGGCCCAACCGGTGCGCTTACCGCCACGAAATCCAGCGGACACTTCTCCAACTG 107739
|
|
|
QY 2581 CGTGACAGGCCACCGCACTGGGCTTGTCTGATCGCTTCGAGCTCATCGAAGACAACTAC 2640
|
|
|
Db 107740 CGTGACAGGCCACCGCACTGGGCTTGTGATCGCTTCGAACTCATCGAAGACAACTAC 107799
|
|
|
QY 2641 GCAGCGCTTAATGATGCTGGGAGCGCCACCAAGCTCACCCATCTCCAAAGTTGTT 2700
|
|
|
Db 107800 GCAGCGCTTAATGATGCTGGGAGCGCCACCAAGCTCACCCATCTCCAAAGTTGTT 107859
|
|
|
QY 2701 GCGCACTTCGCACTCCACTGGTGGTGGCGGTGTAGATCCAGCAGACTTTGCTGCAAGAC 2760
|
|
|
Db 107860 GCGCACTTCGCACTCCACTGGTGGTGGCGGTGTAGATCCAGCAGACTTTGCTGCAAGAC 107919
|
|
|
QY 2761 CCACAAAGTACGATCCCAAGACTCTGTATCGCTTCTTCGCGCGGAGCTTTGGTAAC 2820
|
|
|
Db 107920 CCACAAAGTACGATCCCAAGACTCTGTATCGCTTCTTCGCGCGGAGCTTTGGTAAC 107979
|
|
|
QY 2821 CCTCCAGGTGGCTGGCAGAACCACTGGGACCCCGGCACTGGAGGCGCTCCGAAGGC 2880
|
|
|
Db 107980 CCTCCAGGTGGCTGGCAGAACCACTGGGACCCCGGCACTGGAGGCGCTCCGAAGGC 108039
|
|
|
QY 2881 AAGCACCTCTGACGGAAGTTCTTGAGGAAGACAGCAGCGCACTCGAGCTGATGATCC 2940
|
|
|
Db 108040 AAGCACCTCTGACGGAAGTTCTTGAGGAAGACAGCAGCGCACTCGAGCTGATGATCC 108099
|
|
|

QY 2941 AAGGAACGTGCAACAGCCTCAACCGCCTGCTGTTCCGAAGCCAAACCGAAGATTCTCTC 3000
|
|
|
Db 108100 AAGGAACGTGCAATAGCCTCAACCGCCTGCTGTTCCGAAGCCAAACCGAAGATTCTCTC 108159
|
|
|
QY 3001 GAGCACCGTCCGCGCTTCGGCAACACACTCTGCGCTGATGATCGTGAATCTTCTACGGA 3060
|
|
|
Db 108160 GAGCACCGTCCGCGCTTCGGCAACACACTCTGCGCTGATGATCGTGAATCTTCTACGGA 108219
|
|
|
QY 3061 CTGTCGAGGCGCGGAGACTTTGATCCGCTGCGCAGATGTGCGCACCCCACTGCTTGT 3120
|
|
|
Db 108220 CTGTCGAGGCGCGGAGACTTTGATCCGCTGCGCAGATGTGCGCACCCCACTGCTTGT 108279
|
|
|
QY 3121 CGCTCGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGCGCAACGTC 3180
|
|
|
Db 108280 CGCTCGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGCGCAACGTC 108339
|
|
|
QY 3181 AACGGCCAGATCCGCCCAATGCGTGTGCGTACCGCTCCGTTGAGTCTGTCAACCGCAACC 3240
|
|
|
Db 108340 AACGGCCAGATCCGCCCAATGCGTGTGCGTACCGCTCCGTTGAGTCTGTCAACCGCAACC 108399
|
|
|
QY 3241 GCAGAAAAGGCAGATTCTCCAAACAGGGCCATGTTGCTGCACCAATTCGCTGGTGTGTC 3300
|
|
|
Db 108400 GCAGAAAAGGCAGATTCTCCAAACAGGGCCATGTTGCTGCACCAATTCGCTGGTGTGTC 108459
|
|
|
QY 3301 ACTGTGACTGTTGCTGAAGTGATGAGTCAAGGCTGGAGATGCGATCGCAATCATCGAG 3360
|
|
|
Db 108460 ACCGTGACTGTTGCTGAAGTGATGAGTCAAGGCTGGAGATGCGATCGCAATCATCGAG 108519
|
|
|
QY 3361 GCTATGAAGTGAAGCAACAATCCTGCTTCTGTTGACGCGAAGATTGAACGGTGTG 3420
|
|
|
Db 108520 GCTATGAAGTGAAGCAACAATCCTGCTTCTGTTGACGCGAAGATTGAACGGTGTG 108579
|
|
|
QY 3421 GTTCTCTGCTCAACGAAAGTGAAGGTGGCGACTTGATCGTCTGTTTCTCTAA 3474
|
|
|
Db 108580 GTTCTCTGCTCAACGAAAGTGAAGGTGGCGACTTGATCGTCTGTTTCTCTAA 108633
|
|
|

RESULT 5

AAAX24102
ID AAX24102 standard; DNA; 3728 BP.

XX AAX24102;

XX AC AAX24102;
XX XX
XX 30-JUN-1999 (first entry)

XX C. glutamicum pyruvate carboxylase genomic DNA.

XX KW Pyruvate carboxylase; amino acid production; lysine production;
XX KW threonine production; homoserine production; glutamate production;
XX KW arginine production; feed additive; condiment; pharmaceutical;
XX KW fine chemical; ss.

XX OS Corynebacterium glutamicum.

XX XX Key Location/Qualifiers

XX FH 165..3587

XX FT /*tag= a

XX FT /product= "pyruvate carboxylase"

XX PN DE19831609-Al.

XX XX

XX PD 15-APR-1999.

XX XX PF 14-JUL-1998; 98DE-1031609.

XX XX PR 04-OCT-1997; 97DE-1043894.

XX XX PA (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.

XX XX PI Eikmanns B, Peters-Wendisch P, Sahm H;

XX XX WPI; 1999-245521/21.

XX XX P-PSDB; AAW93971.

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

xx
PT Increasing microbial production of specific amino acids by
PT increasing activity or expression of pyruvate carboxylase
PS Claim 16; Page 8-11; 18pp; German.
XX
XX This invention describes the isolation of a pyruvate carboxylase
CC from Corynebacterium glutamicum which is used in a novel method for
CC production of lysine, threonine, homoserine, glutamate and/or arginine,
CC variously useful as feed additives, condiments, pharmaceuticals and
CC intermediates for fine chemicals. Increasing pyruvate carboxylase
CC activity increases the yield of microbial production of amino acids
CC in the aspartate and/or glutamate families, e.g. about 50% more lysine,
CC 40% more threonine and 150% more homoserine are secreted into the
CC culture medium.
SQ Sequence 3728 BP; 813 A; 1059 C; 1035 G; 821 T; 0 other;
Query Match 97.8%; Score 3397.2; DB 20; Length 3728;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3426; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
Qy 1 GTGACTGCTATCACCTTGGCGGCTCTTTGTTGAAGGAATAATTACTCTAGTGTGCACT 60
Db 114 GTGACTGCTATCACCTTGGCGGCTCTTTGTTGAAGGAATAATTACTCTAGTGTGCACT 173
Qy 61 CACACATCTTCAACGGCTTCAGCATTCAAAAGATCTTTGGTAGCAAAACCGCGGCAATC 120
Db 174 CACACATCTTCAACGGCTTCAGCATTCAAAAGATCTTTGGTAGCAAAACCGCGGCAATC 233
Qy 121 GCGGTCGCTGCTTTCCGTCGACACTCGAAACCGGTGAGCCACCGGTAGCTATTACCCC 180
Db 234 GCGGTCGCTGCTTTCCGTCGACACTCGAAACCGGTGAGCCACCGGTAGCTATTACCCC 293
Qy 181 CGTGAAGATCGGGGATCAATTCACCGCTCTTTGTTGCTTGAAGCTGTCGCGATGGTACT 240
Db 294 CGTGAAGATCGGGGATCAATTCACCGCTCTTTGTTGCTTGAAGCTGTCGCGATGGTACT 353
Qy 241 GAAGGCTCACAGTCAAGCGTACCTGGACATCGATGAATTTATCGGTGAGCTGAAAAA 300
Db 354 GAAGGCTCACAGTCAAGCGTACCTGGACATCGATGAATTTATCGGTGAGCTGAAAAA 413
Qy 301 GTTAAAGCAGATCTATTTACCGGGATATGGCTTCTGTCGTAATAATGCCAGCTTGC 360
Db 414 GTTAAAGCAGATCTATTTACCGGGATACGGCTTCTGTCGTAATAATGCCAGCTTGC 473
Qy 361 CGCGAGTGCAGAAACGSCATTACTTTATTGGCCCAACCCAGAGGTTCTTGATCTC 420
Db 474 CGCGAGTGCAGAAACGSCATTACTTTATTGGCCCAACCCAGAGGTTCTTGATCTC 533
Qy 421 ACCGCTGATAGTCTCGTCGGTAAACCGCGAGAGAGGCTGCTGCGAGTTTGGCG 480
Db 534 ACCGCTGATAGTCTCGTCGGTAAACCGCGAGAGAGGCTGCTGCGAGTTTGGCG 593
Qy 481 GAATCCACCCGAGCAAAACATCGATGAGATCGTTAAAGCGCTGAAGCCAGACTTAC 540
Db 594 GAATCCACCCGAGCAAAACATCGATGAGATCGTTAAAGCGCTGAAGCCAGACTTAC 653
Qy 541 CCNATCTTTGTAAGCAGTTTCGGTGGTGGCGAGCGGATGCGGTTGTTCTTCA 600
Db 654 CCNATCTTTGTAAGCAGTTTCGGTGGTGGCGAGCGGATGCGGTTGTTCTTCA 713
Qy 601 CTTGATGAGCTCCGAAATTTGGCAACAGATCTCTGTAAGCTGAAGCGGCAATTCGG 660
Db 714 CTTGATGAGCTCCGAAATTTAGCAACAGAGCATCTCTGTAAGCTGAAGCGGCTTCGG 773
Qy 661 GACGTTTCGGTATATGTCGAAACGTGCTGATTAACCCCAAGCAATTGAAGTGCAGATC 720
Db 774 GATGGCGGATATATGTCGAAACGTGCTGATTAACCCCAAGCAATTGAAGTGCAGATC 833
Qy 721 CTTGGCGATCGCACTGGAGAGTTGACCTTTATGAACCTGATCTCATCTGACGCT 780
Db 834 CTTGGCGATCACACTGGAGAGTTGACCTTTATGAACCTGATCTCATCTGACGCT 893

Qy 781 CGTCAACAAAAGTTGTCGAAATTCGCGCAGCAGCAGCATTTGGATCCAGAACTGCGTGAT 840
Db 894 CGTCAACAAAAGTTGTCGAAATTCGCGCAGCAGCAGCATTTGGATCCAGAACTGCGTGAT 953
Qy 841 CGCATTTGTGCGGATGCACTAAAGTTCTGCGCTCCATTGGTTTACCAAGGCGCGGGAACC 900
Db 954 CGCATTTGTGCGGATGCACTAAAGTTCTGCGCTCCATTGGTTTACCAAGGCGCGGGAACC 1013
Qy 901 GTGGAATTTCTTGGTCGATGAAAAGGCAACACAGCTTTTTCATCGAAATGAACCCACGTATC 960
Db 1014 GTGGAATTTCTTGGTCGATGAAAAGGCAACACAGCTTTTTCATCGAAATGAACCCACGTATC 1073
Qy 961 CAGGTTGAGCAGCAGCTGAGTCAAGAAGTTCACCGAGGTGAGCTTGTGMAAGGCGCAGATG 1020
Db 1074 CAGGTTGAGCAGCAGCTGAGTCAAGAAGTTCACCGAGGTGAGCTTGTGMAAGGCGCAGATG 1133
Qy 1021 CGCTCGCTGCTGGTGCACCTTGAAGGAATTTGGTCTGACCCAGATATAGATCAAGACC 1080
Db 1134 CGCTCGCTGCTGGTGCACCTTGAAGGAATTTGGTCTGACCCAGATATAGATCAAGACC 1193
Qy 1081 CAGGTTGAGCAGCTGAGTGCAGTCCAGCAGCAGGAGATCCAAACAGGCTTCGCGCCA 1140
Db 1194 CAGGTTGAGCAGCTGAGTGCAGTCCAGCAGCAGGAGATCCAAACAGGCTTCGCGCCA 1253
Qy 1141 GATACCGGAATATATACCGGCTACCGCTCAACAGCGGAGCTGCGCTTCTTTCGAGCT 1200
Db 1254 GATACCGGAATATATACCGGCTACCGCTCAACAGCGGAGCTGCGCTTCTTTCGAGCT 1313
Qy 1201 GCAGCTCAGCTCGGTGGCGAAATCAACCGCACAATTTGATCCATGCTGGTGAATGACC 1260
Db 1314 GCAGCTCAGCTCGGTGGCGAAATCAACCGCACAATTTGATCCATGCTGGTGAATGACC 1373
Qy 1261 TGCGGTGTTTCGACCTTTGAAAATCTGCTGTGTCGTGACAGCGCGCTTGGCTGAGTTTC 1320
Db 1374 TGCGGTGTTTCGACCTTTGAAAATCTGCTGTGTCGTGACAGCGCGCTTGGCTGAGTTTC 1433
Qy 1321 ACCGTGCTGGTGTTCACCAACCAATGCTTTCTTTCGCTGCTGCTCGGGAAGAGGAC 1380
Db 1434 ACCGTGCTGGTGTTCACCAACCAATGCTTTCTTTCGCTGCTGCTCGGGAAGAGGAC 1493
Qy 1381 TTCACTTCAAGCGCATCGCCACCGGATTTATCGCGCATCACCCACACCTCTCTTCAGGCT 1440
Db 1494 TTCACTTCAAGCGCATCGCCACCGGATTTATCGCGCATCACCCACACCTCTCTTCAGGCT 1553
Qy 1441 CCACCTGCGGATGATGAGCAGGAGCGCATCTCTGGATTACTTTGGCAGATGTCAACGCGAAC 1500
Db 1554 CCACCTGCTGATGATGAGCAGGAGCGCATCTCTGGATTACTTTGGCAGATGTCAACGCGAAC 1613
Qy 1501 AAGCTCATGTTGCTGCTCCAAAGGATGTTGAGCAGCAACCAATCGATTAAGCTGCCCAATC 1560
Db 1614 AAGCTCATGTTGCTGCTCCAAAGGATGTTGAGCAGCTCTATCGATTAAGCTGCCCAATC 1673
Qy 1561 AAGGATCTGCCACTGCCACCGCTTCCGTCGACCGCTGACCGCTGAGAGCTTGGCCAGCCGC 1620
Db 1674 AAGGATCTGCCACTGCCACCGCTTCCGTCGACCGCTGAGAGCTTGGCCAGCCGC 1733
Qy 1621 TTTGCTCGTATCTCGTCAGCAGGAGCGCAGCTTACTGATACCACTTCCGCGAT 1680
Db 1734 TTTGCTCGTATCTCGTCAGCAGGAGCGCAGCTTACTGATACCACTTCCGCGAT 1793
Qy 1681 GCACACAGCTTTTGGTGGACCGGCTCGCTCATTTTCGCACTGAAGCTTCGCGCAGAG 1740
Db 1794 GCACACAGCTTTTGGTGGACCGGCTCGCTCATTTTCGCACTGAAGCTTCGCGCAGAG 1853
Qy 1741 GCGTTCGCAAGCTGATCTCGTGAAGCTTTTTCGTCGAGGCTTCGGGCGCGGCGCTAC 1800
Db 1854 GCGTTCGCAAGCTGATCTCGTGAAGCTTTTTCGTCGAGGCTTCGGGCGCGGCGCTAC 1913
Qy 1801 GATGTGCGATGCTGTTTCTTTTGAAGATCCGTGGGACAGGCTCGAGCTCGCGGAG 1860
Db 1914 GATGTGCGATGCTGTTTCTTTTGAAGATCCGTGGGACAGGCTCGAGCTCGCGGAG 1973

Qy	1861	CGGATGCCGAATGTAAACATTCAAGATGCTGCTTTCGCGCGCCGAAACACCGTGGGATACACC	1921
Db	1974	CGGATGCCGAATGTAAACATTCAAGATGCTGCTTTCGCGCGCCGAAACACCGTGGGATACACC	2033
Qy	1921	CCGTACCCAGACTCCGTCCTGCGCGCGTTCGTTTAAAGGAAGCTGCCAGCTCCGGCGTGCAC	1980
Db	2034	CCGTACCCAGACTCCGTCCTGCGCGCGTTCGTTTAAAGGAAGCTGCCAGCTCCGGCGTGCAC	2093
Qy	1981	ATCTTCCGCGCTCTTCGACGCGCTTAAACGACGCTCTCCAGATGCGTCCAGCAATCGACGCA	2040
Db	2094	ATCTTCCGCGCTCTTCGACGCGCTTAAACGACGCTCTCCAGATGCGTCCAGCAATCGACGCA	2153
Qy	2041	GTCTCGAGAGCCAAACACCGCGGTAGCCGAGGTGGCTATGCGTTATTCCTGGTGATCTCTCT	2100
Db	2154	GTCTCGAGAGCCAAACACCGCGGTAGCCGAGGTGGCTATGCGTTATTCCTGGTGATCTCTCT	2213
Qy	2101	GATCCAAATGAAAGCTCTACACCCCTGGATTAACCTTAAGATGCGCAGAGGAGATCGTC	2160
Db	2214	GATCCAAATGAAAGCTCTACACCCCTGGATTAACCTTAAGATGCGCAGAGGAGATCGTC	2273
Qy	2161	AAGTCTGGCGCTCACATCTTCCTGGGCCATTAAGGATATAGGCTGGTCTCTTCGCGCCAGCTGCG	2220
Db	2274	AAGTCTGGCGCTCACATCTTCGGGCCATTAAGGATATAGGCTGGTCTCTTCGCGCCAGCTGCG	2333
Qy	2221	GTAAACCAAGCTGCTACCGGCACTGCGCGCTGAATTCGATCTGCCAGTGACGTGCACACC	2280
Db	2334	GTAAACCAAGCTGCTACCGGCACTGCGCGCTGAATTCGATCTGCCAGTGACGTGCACACC	2393
Qy	2281	CAGACACTCGGGGTGGCCAGTTGGCTACTACTTGTCTGCGAGCTCAAGCTGTGGAGAT	2340
Db	2394	CAGACACTCGGGGTGGCCAGCTTGGCAACCTTACTTGTCTGCGAGCTCAAGCTGTGGAGAT	2453
Qy	2341	GCTGTTGACGGTGTTCGCGACACACATGTCCTGGCAACAACCTCCAGGCCATCCCTCTCTGCC	2400
Db	2454	GCTGTTGACGGTGTTCGCGACACAATGTCCTGGCAACAACCTCCAGGCCATCCCTCTCTGCC	2513
Qy	2401	ATTGTTGCTGCATTCGGGCAACACCGTTCGCGAATACCGGTTTGAAGCTCGAGGCTGTTTCT	2460
Db	2514	ATTGTTGCTGCATTCGGGCAACACCGTTCGCGAATACCGGTTTGAAGCTCGAGGCTGTTTCT	2573
Qy	2461	GACCTCGAGCCGTACTCGGGAAAGCTGTGCGGCACTGTACTCTGCCATTTGAGCTCGAGGCTGTTTCT	2520
Db	2574	GACCTCGAGCCGTACTCGGGAAAGCTGTGCGGCACTGTACTCTGCCATTTGAGCTCGAGGCTGTTTCT	2633
Qy	2521	CCAGGCCCAACCGGTTCGCGTCTACGCCACAGAAATCCACAGCGGACAGTTGCTGCCAAACCTG	2580
Db	2634	CCAGGCCCAACCGGTTCGCGTCTACGCCACAGAAATCCACAGCGGACAGTTGCTGCCAAACCTG	2693
Qy	2581	CGTGCAAGGCCACCGCACTCGGCGCTTGTGATCGCTTCGAGCTCATCGAAGACAACTAC	2640
Db	2694	CGTGCAAGGCCACCGCACTCGGCGCTTGTGGATCGTTTTCGAATCATCGAAGACAACTAC	2753
Qy	2641	GCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCAACCCATCTCTCAAGTTGTT	2700
Db	2754	GCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCAACCCATCTCTCAAGTTGTT	2813
Qy	2701	GGCGACTCTCGACTCCACCTCGTTTGGTGGGTGTAGATTCAGACAGACTTTTGTCTGCAGAC	2760
Db	2814	GGCGACTCTCGACTCCACCTCGTTTGGTGGGTGTAGATTCAGACAGACTTTTGTCTGCAGAC	2873
Qy	2761	CCACAAAGTACGACATCCAGACTCTGTCACTCGGTTTCTCTCGCGCGGCGAGCTTGGTAAAC	2820
Db	2874	CCACAAAGTACGACATCCAGACTCTGTCACTCGGTTTCTCTCGCGCGGCGAGCTTGGTAAAC	2933
Qy	2821	CCTTCAGGTGGCTGGCCAGAAACCATGCGCAACCCCGCACTGGGAAGGCGCGCTCCGAAGGC	2880
Db	2934	CCTTCAGGTGGCTGGCCAGAGCCATCGCGCAACCCCGCACTGGGAAGGCGCGCTCCGAAGGC	2993
Qy	2881	AAGSCACTCTGACGGAAGTTCTCTGAGGAGAGCAGGCGCACTCTCGAGCGTGTGATTC	2940
Db	2994	AAGSCACTCTGACGGAAGTTCTCTGAGGAGAGCAGGCGCACTCTCGAGCGTGTGATTC	3053
Qy	2941	AAGGAACGTCGCAACACGCTCAACCGCTGCTGTTTCCCGAAGCCAAACGGAAGACTTCCTC	3000

Db	3054	ANGGAAGTGCATATGCTCAACCCCTGCTGTTCCTCCGAAGCCAACGAAGAGTTCCTC	3113
Qy	3001	GAGCACCGTCGCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTTCTTACGGA	3060
Db	3114	GAGCACCGTCGCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTTCTTACGCG	3173
Qy	3061	CTGTGTCAGGGCCGCGAGACTTTGATTCGCCCTGCGCAGATGTGCGCACCCCACTGCTTGT	3120
Db	3174	CTGTGTCAGGGCCGCGAGACTTTGATTCGCCCTGCGCAGATGTGCGCACCCCACTGCTTGT	3233
Qy	3121	CGCTGTGATGCGATCTCTGAGCCAGACGATAAAGGGTATGCGCAATGTTGTGGCCAACGTC	3180
Db	3234	CGCTGTGATGCGATCTCTGAGCCAGACGATAAAGGGTATGCGCAATGTTGTGGCCAACGTC	3293
Qy	3181	AACGGCCAGATTCGCCCAATGCGTGTGCTGACCGCTCCGTTGAGTCTGTCAACGCAACC	3240
Db	3294	AACGGCCAGATTCGCCCAATGCGTGTGCTGACCGCTCCGTTGAGTCTGTCAACGCAACC	3353
Qy	3241	GCAGAAAGGCGAGATTCCTCCAAAGGCCCATGTTGTGCAACCATTCGCTGGTGTGTC	3300
Db	3354	GCAGAAAGGCGAGATTCCTCCAAAGGCCCATGTTGTGCAACCATTCGCTGGTGTGTC	3413
Qy	3301	ACTGTGACTCTTCTGAGGTGATGAGGTCAAGGCTGAGATGCAATGCGCAATCATCGAG	3360
Db	3414	ACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGAGATGCAATGCGCAATCATCGAG	3473
Qy	3361	GCTATGAAGATGGAAGCAACATCACTGCTTCTGTTGACGCAAGATTGAACGCTTGTG	3420
Db	3474	GCTATGAAGATGGAAGCAACATCACTGCTTCTGTTGACGCAAGATTGAACGCTTGTG	3533
Qy	3421	GTTCCTCTGCAACGAAGGTGGAAGTGGCGACTTGTGATCGTCGCTTTCCTAA	3474
Db	3534	GTTCCTCTGCAACGAAGGTGGAAGTGGCGACTTGTGATCGTCGCTTTCCTAA	3587
RESULT 6			
AAH65730	ID	AAH65730 standard; DNA; 3420 BP.	
XX	AC	AAH65730;	
XX	DT	26-SEP-2001 (first entry)	
XX	DE	C glutamicum coding sequence fragment SEQ ID NO: 765.	
XX	KW	Corynebacterium; amino acid synthesis; vitamin; saccharide;	
XX	KW	organic acid synthesis; ds.	
XX	OS	Corynebacterium glutamicum.	
XX	PN	EPI108790-A2.	
XX	PD	20-JUN-2001.	
XX	PF	18-DEC-2000; 2000EP-0127688.	
XX	PR	16-DEC-1999; 99JP-0377484.	
PR	PR	07-APR-2000; 2000JP-0159162.	
PR	PR	03-AUG-2000; 2000JP-0280988.	
XX	XX	(KYOW) KYOWA HAKKO KOGYO KK.	
XX	PI	Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;	
PI	PI	Tateishi N, Senoh A, Ikeda M, Ozaki A;	
XX	XX	WPI; 2001-376931/40.	
DR	DR	P-PSDB; AAG90511.	
XX	XX	Novel polynucleotides derived from Corynebacterium bacteria, for identifying	
PT	PT	mutation point of a gene, measuring expression of a gene, analysing	
PT	PT	expression profile or pattern of a gene and identifying homologous gene	

Qy 1912 GGATACACCCCGTACCCAGACTCCGTCGCGCGCGGTTTGTAAAGAAAGCTGCCAGCTCC 1971
Db 1861 GGATACACCCCGTACCCAGACTCCGTCGCGCGCGGTTTGTAAAGAAAGCTGCCAGCTCC 1920
Qy 1972 GCGGTGGACATCTTCCGATCTTCAGCGCGTTAAACGACGTCCTCCAGATGCGTCCAGCA 2031
Db 1921 GCGGTGGACATCTTCCGATCTTCAGCGCGTTAAACGACGTCCTCCAGATGCGTCCAGCA 1980
Qy 2032 ATCGACGAGTCTCGGAGACCAACACCGCGGTAGCCGAGTGGCTATGCTTATCTCGT 2091
Db 1981 ATCGACGAGTCTCGGAGACCAACACCGCGGTAGCCGAGTGGCTATGCTTATCTCGT 2040
Qy 2092 GATCTCTGATCAAAATGAAGCTCTACACCGCGGTCTACCTTAAAGATGGCAGAG 2151
Db 2041 GATCTCTGATCAAAATGAAGCTCTACACCGCGGTCTACCTTAAAGATGGCAGAG 2100
Qy 2152 GAGATCGTCAAGTCTGGCGCTCACATCTGGCCATTAAAGGATATGGCTGCTCTCGC 2211
Db 2101 GAGATCGTCAAGTCTGGCGCTCACATCTGGCCATTAAAGGATATGGCTGCTCTCGC 2160
Qy 2212 CCAGCTCGGTTAAACAAAGCTGGTCAACCGCACTGGCGCGTGAATTCGATCTGCCAGTGCAC 2271
Db 2161 CCAGCTCGGTTAAACAAAGCTGGTCAACCGCACTGGCGCGTGAATTCGATCTGCCAGTGCAC 2220
Qy 2272 GTGCACACCCAGACACTGCGGTTGGCAGTTGGCTACCTACTTTGCTGCAAGCTCAAGCT 2331
Db 2221 GTGCACACCCAGACACTGCGGTTGGCAGTTGGCTACCTACTTTGCTGCAAGCTCAAGCT 2280
Qy 2332 GGTCCAGATGCTGTTGAGGCTGTTCCGACCACTGCTGCGCACCACTCCCGACCACTCC 2391
Db 2281 GGTCCAGATGCTGTTGAGGCTGTTCCGACCACTGCTGCGCACCACTCCCGACCACTCC 2340
Qy 2392 CTGTCGCAATGTTGTGCAATTCGCGCACACCGCTGCGCATACCGGTTTGAGCTCGAG 2451
Db 2341 CTGTCGCAATGTTGTGCAATTCGCGCACACCGCTGCGCATACCGGTTTGAGCTCGAG 2400
Qy 2452 GCTGTTTCTGACCTCGAGCCGTAATGGAAGCTGTGGCGGAGCTGTACCTGCCATTTGAG 2511
Db 2401 GCTGTTTCTGACCTCGAGCCGTAATGGAAGCTGTGGCGGAGCTGTACCTGCCATTTGAG 2460
Qy 2512 TCTGGAACCCAGGCCCAACCGGTCGGTCTACCGCACGAAATCCAGGCGGACAGTTG 2571
Db 2461 TCTGGAACCCAGGCCCAACCGGTCGGTCTACCGCACGAAATCCAGGCGGACAGTTG 2520
Qy 2572 TCCAACTCTGCTGCACAGGCCACCGCACTGGCGCTTGTGATCGCTTCGAGCTCATCGAA 2631
Db 2521 TCCAACTCTGCTGCACAGGCCACCGCACTGGCGCTTGGGATCGTTTCCAACTCATCGAA 2580
Qy 2632 GACAACTACGAGCCGTTAATGAGATGCTGGAGCGCCCAACCAAGGTCAACCCATCTCTCC 2691
Db 2581 GACAACTACGAGCCGTTAATGAGATGCTGGAGCGCCCAACCAAGGTCAACCCATCTCTCC 2640
Qy 2692 AAGGTTGTTGGCGACCTCGCACTCCACTGCTGTTGGTGGGTGTAGATCCAGCAGACTTT 2751
Db 2641 AAGGTTGTTGGCGACCTCGCACTCCACTGCTGTTGGTGGGTGTAGATCCAGCAGACTTT 2700
Qy 2752 GCTCAGACCCACAAAAGTACGACATCCGACACTCTGTATCGGCTTCTCGCGGGCGAG 2811
Db 2701 GCTCAGACCCACAAAAGTACGACATCCGACACTCTGTATCGGCTTCTCGCGGGCGAG 2760
Qy 2812 CTTGGTAACCTTCAGGTGGCTGGCCAGAACCACTGCGCACCCCGCGCACTGGAAAGGCCGC 2871
Db 2761 CTTGGTAACCTTCAGGTGGCTGGCCAGAACCACTGCGCACCCCGCGCACTGGAAAGGCCGC 2820
Qy 2872 TCCGAAGCAAGGCACTCTGACGGAAGTCTCTGAGGAAGAGGCGCACTCGAGCT 2931
Db 2821 TCCGAAGCAAGGCACTCTGACGGAAGTCTCTGAGGAAGAGGCGCACTCGAGCT 2880
Qy 2932 GATGATTCCAGGAAGCTGCGAACAGCTCAACCGCTGCTGTTCCCGAAGCCCAACCGAA 2991
Db 2881 GATGATTCCAGGAAGCTGCGAACAGCTCAACCGCTGCTGTTCCCGAAGCCCAACCGAA 2940

Qy 2992 GAGTTCTCGAGCACCGTCCGCGCTTCGGCAACACCTCTCGGCTGGATGATCGTGAATTC 3051
Db 2941 GAGTTCTCGAGCACCGTCCGCGCTTCGGCAACACCTCTCGGCTGGATGATCGTGAATTC 3000
Qy 3052 TTCTACGGACTGGTTCGAGGCGCGGAGACTTTGATCCGCTGCCAGATGTGCGCACCCCA 3111
Db 3001 TTCTACGGACTGGTTCGAGGCGCGGAGACTTTGATCCGCTGCCAGATGTGCGCACCCCA 3060
Qy 3112 CTGCTTCTCCCTCGGATGCGATCTCTGAGCCAGACGATAAAGGATATGCGCAATGTTGTG 3171
Db 3061 CTGCTTCTCCCTCGGATGCGATCTCTGAGCCAGACGATAAGGATATGCGCAATGTTGTG 3120
Qy 3172 GCCAAGCTCAACGGCCAGATCCGCCCAATCGGTGTGCGTGCACCGCTCCGTTGAGTCTGTC 3231
Db 3121 GCCAAGCTCAACGGCCAGATCCGCCCAATCGGTGTGCGTGCACCGCTCCGTTGAGTCTGTC 3180
Qy 3232 ACCGCAACCCAGAAAAGGAGGATTCCTCCAAAGGGCCATGTTGTCGACCAATTCGCT 3291
Db 3181 ACCGCAACCCAGAAAAGGAGGATTCCTCCAAAGGGCCATGTTGTCGACCAATTCGCT 3240
Qy 3292 GGTGTTGCTCACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCGAGTCGCA 3351
Db 3241 GGTGTTGCTCACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCGAGTCGCA 3300
Qy 3352 ATCATCGAGCTATGAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAGATTGAA 3411
Db 3301 ATCATCGAGCTATGAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAAATCGAT 3360
Qy 3412 CGGTTGCTGCTTCTGCTGCAACGAGGTGGAAGGTGCGGACTTTGATCGCTGCTGTTTCC 3471
Db 3361 CGGTTGCTGCTTCTGCTGCAACGAGGTGGAAGGTGCGGACTTTGATCGCTGCTGTTTCC 3420

RESULT 7
AAF87437
ID AAF87437 standard; DNA; 4013 BP.
XX
AC AAF87437;
XX
DT 09-JUL-2001 (first entry)
XX
DE Corynebacterium thermoaminogenes pc nucleotide sequence.
XX
KW Corynebacterium; thermophilic; amino acid biosynthesis; enzyme;
KW thermotolerant; aceA; accBC; dtsR1; dtsR2; pfk; scrB; gluABCD;
KW pdhA; pc; ppc; acn; icd; lpd; odhA; ds.
XX
OS Corynebacterium thermoaminogenes.
XX
Key Location/Qualifiers
FH 319..3738
FT CDS /*tag= a
FT /product= "pc protein"
XX
PN WO200125447-A1.
XX
PD 12-APR-2001.
XX
PF 04-OCT-2000; 2000WO-JP06913.
XX
PR 04-OCT-1999; 99JP-0282716.
PR 01-NOV-1999; 99JP-0311147.
PR 21-APR-2000; 2000JP-0120687.
XX
PA (AJIN) AJINOMOTO CO INC.
XX
PI Hirano S, Nonaka G, Matsuzaki Y, Akiyoshi N, Nakamura K, Kimura E;
PI Osumi T, Matsui K, Kawahara Y, Kurahashi O, Nakamatsu T;
PI Sugimoto S;
XX
DR WPI; 2001-300170/31.
XX
P-PSDB; AAB83180.

PT Proteins and their DNA useful for microbial production of L-amino acids

XX -

PS Claim 34; Page 126-132; 215pp; Japanese.

XX The present sequence is provided in a specification relating to genes
CC encoding thermophilic amino acid biosynthesis system enzymes of
CC the thermotolerant bacterium *Corynebacterium thermoaminogenes*.
CC The novel proteins retain at least 30% isocitrate lyase activity
CC after heating at 500C for 5 minutes. DNA fragments encoding the
CC enzymes were isolated from a *Corynebacterium thermoaminogenes*
CC chromosomal DNA plasmid library by PCR. The DNA may be used for
CC developing strains of amino acid producing microorganisms.

XX Sequence 4013 BP; 735 A; 1376 C; 1234 G; 668 T; 0 other;

Query Match 64.9%; Score 2256.2; DB 22; Length 4013;
Best Local Similarity 78.7%; Pred. No. 0;
Matches 2734; Conservative 0; Mismatches 733; Indels 8; Gaps 3;

QY 1 GTGACTGCTATCACCC--TTGGCGGTCTCTGTGAAAGGAATAATTACTCTAGTGTGCA 58
DB 269 GTGACTGCCATCACCCACAGGCTGTCTCTGTGAAAGGAACAAAACCTGTGGTTCAA 328
QY 59 CTCACATCTTCAAGCTTCCAGATTCAGAAAGATCTTGTAGCAACCCGCGGAAA 118
DB 329 C---AACACCTTCAGCTGCGCGGTTCAGAAAGATCTGTGGCCACCCAGGTGAAA 385
QY 119 TCGGGTCCGTCTTCCGTGAGACTCGAAACGGTGCAGCAGGTAGTATTTACC 178
DB 386 TCGGGTGCAGGATCTCCGCGCGCTTACGACCGGGCCCAACCGTGGCCATCTACC 445
QY 179 CCCTGGAAGATCGGGATCATCCACCGCTTTTGTCTTCTGAAAGCTGTCCGCTGTTGA 238
DB 446 CCGGGAGGACCGTGGCTCTTCCACCGTCTCTTCCCTCCGAGCGGTGAGGATCGGAA 505
QY 239 CTGAAGGCTCACCAGTCAAGGCTACCTGGACATCGATGAAATATTCGGTGCAGCTAAA 298
DB 506 CCGAGGCTCACCCTCAAGGCTACCTCGATATTTGATGAGATCATCAACCCGCGCAAGA 565
QY 299 AAGTTAAAGCAGATGCTATTATCCCGGATATGGCTTCTCTGTCTGAAATGCCAGCTTG 358
DB 566 AGGTGAAAGCGGACCGGTCTACCCGGGTATGGTTTCTCTTTCGAAATATGCCAGCTCG 625
QY 359 CCGCGAGTGGCGGAAACCGCATTTATTTGCGCCCAACCCAGAGGTTCTTGATC 418
DB 626 CGCTGAAATGCGGGAACCGCATTTCTATCGTCTCCACCCCGAGGTGCTCGACC 685
QY 419 TCACCGTGATAGTCTCGTGGTAAACCGCGGAAGAGCTGTCTGCCAGTTTGG 478
DB 686 TCACGGGACAGTCAAGGCTGTCTCCGCGGAAGAGCGGGCTGCGGTGCTGG 745
QY 479 CGGAATCCACCCGAGCAAAACATCGATGACATCGTTAAAGCGCTGAAGCCAGACTT 538
DB 746 CGGAATCCACCCGAGCAACCGATCGATGAGATCGTCAAGAGTCCGAGGGGCGAGACT 805
QY 539 ACCCCATTTTGAAGGCAAGTTGCGGTGGTGGGGAACCGGTATGCGCTTTGTTCTT 598
DB 806 ACCCGATTTCTGTCAGAGCCGTGCGAGGTGGTGGGGGCTGTATGCGTTTCGCGAGA 865
QY 599 CACTGTAGCTCCGCAATTTGGCAACAGAGCATCTCGTGAAGTGAAGCGGCAITTCG 658
DB 866 AGCCCGAGGACCTGCGTGAGCTGGCAGAGGAGCCCTCCCGAGGCGGAGGCGGTTTCG 925
QY 659 GCGACGGTTCGGTATATGTGCAAGCTGCTGTGATTAACCCCGCAGCACATTTGAAGTGCAGA 718
DB 926 GTGACCGATCCGTCTACGTCCAAACGGGCGCTGTATCAAAACCCAGCACATTCGAGGTGCAGA 985
QY 719 TCCTTGGCGATCGCACTGGAGAGTGTGTACACCTTTATGAAGTGAAGTCACTCACTGCAGC 778
DB 986 TCCTTCGTGATCACACCGGCGATGTATCATCCACCTGTATGAACCGACTGTTCCTTCGACG 1045
QY 779 GTCTGTACCAAAAAGTTGTCGAAATTCGCGCCAGCAGCATTTGGATCCAGAACTGCGTG 838

DB 1046 GCGGCCACCAAGGTCGTGGAGATCGACCTGCCAGCACCTCGACCCGAGCTCGCG 1105
QY 839 ATCGCATTTGTGCGGATGTCAGTAAAGTTCTGCCGCTCCATTGGTTACCAAGGCGCGGAA 898
DB 1106 ACCGATCTGTGCGGATGCGTGAAGTTCTGCAATTCATCGATACCAAGGCGCGCGCA 1165
QY 899 CGGTGGATCTTGGTGGATGAAGGCAACACCGTTTTCATCGAAATGAACCCACGTA 958
DB 1166 CCGTGGATTTCTGTCGACGAGCGGGCAACACGCTTTCATTGAGATGAACCCCGCA 1225
QY 959 TCCAGTTGACACACCGTGTACTGAAAGTACACGAGGTGGACCTGTGAAGGCGCAGA 1018
DB 1226 TCCAGTTGAAACACACCGTGTGACGAGGAGGTCACTCCGTCGACCTGTGTCAGGCGCAGA 1285
QY 1019 TCGCTTCCGTGCTGTGTCACACTTTGAAGGAAATTTGGTCTGACCCCAAGATAAGATCAAGA 1078
DB 1286 TGCACCTTGGCGCGCGGTGCCACCTCGAAGGAACCTGGGCCCTGACCCAGGACACAGATCACCA 1345
QY 1079 CCCAGGTTGCAGCACTGCAAGTGGCGCATCACACGGAAGATCCAAACAAACGGTTCGCGC 1138
DB 1346 CCCAGGTTGCCGCTTGCAGTGGCGCATCACACGAGGACCCGTCACCAACACTTCGCGC 1405
QY 1139 CAGATACCGGAACTATCACCGCTTACCGCTCACAGGCGGAGCTGGCGTTCGCTTTGACG 1198
DB 1406 CCGACACCGGTGTGATCACCGCTTACCGCTCCCGGGTGGTGGGTGTCGCTCGACG 1465
QY 1199 GTGCAAGTCACTGCTGGTGGGGAATCACCGCACACTTTTGACTCCATCTGCTGGTGAATAATGA 1258
DB 1466 GCGCAGCCAGCTCGGCGCGAGATCACCGCACATTTTCGATTCCATGCTGTTCAAGATGA 1525
QY 1259 CCGCGGTGTTCCGACTTTGAACTGCTGTTGCTGTGACAGCGCGGTTCGGTTCGAGT 1318
DB 1526 CTTGCGCGGTTTCGATTTTCGAGACCCGCTGTCCGAGCCGAGCGCCCTCGCGCGAGT 1585
QY 1319 TCACGCTGTGTTGTTGCAACCAACATTTGTTTTCGCTGCGTTCGCTGCGGGAAGAGG 1378
DB 1586 TCACGCTTCCGCGGTGGCCACCAACATTCGCTTTCGCTGCGTTCGCTGCGGGAAG 1645
QY 1379 ACTTCACCTTCAAGCGCATCGCCACCGGATTTATPCGGCGATCACCCACACTCTCTTCAGG 1438
DB 1646 ACTTCACCAAGAGCGCATCGACACCGGCTTTCATPCGGCTCCCAACAGCACTCTCTCAGG 1705
QY 1439 CTCACCTTCGCGATGATGAGCAGGAGCGCATCTCTGGATTTACTTTGGCAGATGTCAACGTGA 1498
DB 1706 CCCCACCGCGCGACGATGAGCAGGCGGCGATCTCTGGAATACCTGGCGGATGTCAACGTGA 1765
QY 1499 ACAAGCTCATGTTGCTGCTCCAAAGGATTTGTCAGCACCAATCGATAAGCTGCCCAACA 1558
DB 1766 ACAAAACCCACCGGTGAACGCCC---CGAGACAGCCCGTCCGATAGAGAAGCTGCGCGAGG 1822
QY 1559 TCAAGGATCTGCCACTGCGCCACGCGTTCCGTCGACCGCTGAGCAGCTTGGCCCGAGCGG 1618
DB 1823 TGGAGAACATCCCGTGCACACGCGCTCCCGGACCGCTGAGCAGCTCGGCCCGAGG 1882
QY 1619 CGTTTGTCTGTGATCTCCGTGAGCAGGACGACCTGGCAGTTACTGATACCACTTCGCGG 1678
DB 1883 GTTTCGCGCGATCTGCGCAACAGATGCCCTTGGCCGTCACCGACACCACTTCGCGG 1942
QY 1679 ATGCAACAGCTTTTGTTCGGAACCGGTCCTCATTTGCGACCTGAAGCTTCGCGCAG 1738
DB 1943 ATGCCACAGTCCCTCTCTGGCCACCCGCTGCGCTCTCTGCGCTGACCCCGCGCGC 2002
QY 1739 AGCCGCTCGCAAGCTGATCTCTGAGCTTTTGTTCGTGGAGGCTTGGGGCGCGCGACCT 1798
DB 2003 GGGCGCTCGCAAGCTCAACCCCGAGCTGCTGCTGGTGGAGGCTTGGGGCGGTGCACCT 2062
QY 1799 AGCATGTGCGATGCTTTTCTCTTTGAGGATCCGTGGGACAGGCTCGACAGCTCGCGG 1858
DB 2063 ACAGGTGCGCATGCGCTTCTCTTCGAGGATCCGTGGGCAACGCTCGATGAGTGGGTG 2122
QY 1859 AGGCGATGCGGAATGTAAACATTCAGATGCTGCTTCGCGGCGCGCAACCGCTGGGATACA 1918

Db 2123 AGCGATGCCGAATCTGAACATCCAGATGCTGCTCGGTGCTCGCAACACCGTCCGGGTACA 2182
Qy 1919 CCCGTAACCCAGACTCCGCTGTCGCGCGGTGTTGAAGAAAGTGCACAGCTCCCGCGTGG 1978
Db 2183 CCCGTAACCCGATTCGGTGTGCGCGCGTGTGTCAGGAGCGCCCAAGTCCCGTGTGG 2242
Qy 1979 ACATCTTCGCGATCTTCGACCGGCTTAACAGCGTCTTCCAGATGGTCCAGCAATCCGACG 2038
Db 2243 ACATCTTCGCGATCTTCGACCGGCTTAACAGCACTCTCCAGATGGCCCGCCATCCGACG 2302
Qy 2039 CAGTCTCTGGAGCAACACACCGCGGTAGCCGAGGTGGCTTATGGCTTATTTCTGGTGTCTCT 2098
Db 2303 CCGTCTCTGGAGCAACCGGCAACAGTGTGCGGAGGTCCGCAATGGCGTACTTCCGGTGACCTGT 2362
Qy 2099 CTGATTCMAATGAAGTCTTACACCTCTGGAATTAACCTAAGATGGCAAGAGATCG 2158
Db 2363 CCAATTCGGGGGGAAGTCTTACACCTCTGGAATTAACCTAAGATGGCAAGAGATCG 2422
Qy 2159 TCAAGTCTGGCGCTCACATCTTGCGCATTAAGGATATGGCTGTGCTTTCGCGCCAGCTG 2218
Db 2423 TCGACTCCGGTGCACACATCTTGCGCATTAAGGATATGGCTGTGCTTTCGCGCGCGCG 2482
Qy 2219 CGGTAAACCAAGCTGTTCACCGCACTGCGCGCTGAATTCGATCTGCGAGTGCACGTGCACA 2278
Db 2483 CGGCGCCCAACTGTTCACCGCTCTGCGCGTGAATTCGATCTGCGCGTGCATGTCCACA 2542
Qy 2279 CCCACGACATCGCGGTGGCGAGTGGCTTACTTCTGCTGCACTCAAGCTCTGCTGCAG 2338
Db 2543 CCCACGACATCGCGGTGGCGAGTGGCTTACTTCTGCTGCACTCAAGCTCTGCTGCAG 2602
Qy 2339 ATGCTGTGAGGTGCTTCCGACCACTGTCTGGCAACCACTCCGAGGCATCCCTGTCTG 2398
Db 2603 ATGCGGTGCAGCGCGCTTCGCACTCCGCTGCTCGGTATCCACCTCCGAGCGGTGATGTCG 2662
Qy 2399 CCATGTTGCTGCTATTCGCGCACACCGCTGCGGATACCGGTTTCGAGCTCGAGGCTGTTT 2458
Db 2663 CTCTGTTGCGCGTGTGCGGACACCGGACGCAACCGGCTCAACCTGAGCGCGTCT 2722
Qy 2459 CTGACCTCGAGCGCTACTGGAAGCTGTGCGCGGACTGTACTCTGCCATTTGAGTCTGGAA 2518
Db 2723 CCGACCTGGAACCGTACTGGAAGCGGTTCGCGGACTGTACTCTGCCGTTTGAATCCGGCA 2782
Qy 2519 CCCAGGCGCAACCGGTGCGTCTACGCGCACGAAATCCGAGGCGGACAGTTGTCCAAAC 2578
Db 2783 CCCGCGCGCGGACCGGCGGCTTTACGCGCACGAGATCCCGCGGCGTCACTGTCCAAAC 2842
Qy 2579 TGCTGTGACAGCGCACCGCACTGCGGCTTGTGATCGCTTCGAGCTCATCAAGCACT 2638
Db 2843 TGCTGTGACAGCGCGTGTGCACTGCGGCTTGTGCGGACCGCTTCGAGCTCATCGAGGACT 2902
Qy 2639 ACGCAGCGGTTAATGAGATGCTGGGAGCGCCCAACCAAGGTCAACCCCATCTCTCCAAAGTTG 2698
Db 2903 ACGGCGCGTCAACGAGATGCTGGGTGCTCGGACCAAGGTCAACCCGCTCTCCAAAGTTG 2962
Qy 2699 TTGCGGACTCGCACTCAACCTGTTGTTGGGTGTGAGATCCAGCACTTTGCTGAG 2758
Db 2963 TCGGTGACCTCGCACTGCACTCTGCTGCGGTGCGGTGTGAGCCCGGAGGATTTCCGCGCG 3022
Qy 2759 ACCCACAAGTACGATCCAGCACTGTGATCGGCTTCTGCGCGGCGAGCTTGTA 2818
Db 3023 ATCCGAGAGTACGATCCCGATTCGGGTATCGGCTTCTCCGCGGCGAATCGGGTA 3082
Qy 2819 ACCCTCAGGTGGTGGCGAGAACCACTGCGCACCCCGGCACTGGAAGGCGGCTCCGAAG 2878
Db 3083 CCCCTCCCGTGGTGGCGGAGAACCGGTGCGCACCCGTCGACTCGAGGTCGCTCCAGG 3142
Qy 2879 GCAAGGCACTCTGACGGGAAGTTCTTGAAGAGAGCGGCGACCTCGAGCTGATGATT 2938
Db 3143 GTAAGGCGCGCTGGCGAGATCCCGCGAGGAGCGGCGCACTTGAATTCGATGATT 3202
Qy 2939 CCAAGGAACCTCGCAACAGCTGTCAACCGCTGTGTTCCGAGCGCAACCGAGAGTTC 2998
Db 3203 CCGCGGAGCGTTCGCGGCAACCTCAACCGCTGTGTTCCGAGCGGAGGAGTTC 3262

Qy 2999 TCGAGCAACCGTTCGCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTCCTCTACG 3058
Db 3263 TTGAGCAACCGTTCGCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTCCTCTACG 3322
Qy 3059 GACTGTGCGAGGCGCGAGACTTTGATCGCCTGCGCAGATGTCGGCACCCATCGCTTG 3118
Db 3323 GCTTGAAGGAGGAGCGTGAAGAGCTGATCCGACTGACCGGTGTGTCCACCCGATGGTGG 3382
Qy 3119 TTCCCTCTGGATGCGATCTCTGAGCCAGACGATAGGGTATGCGCAATGTTTGTGGCCAACG 3178
Db 3383 TCCGCTCTGGATGCGGTGTCGAAACCGGATGACAAAGGATGCGCAACGTTGTTGTCACG 3442
Qy 3179 TCAACGCGCAGATCCGCGCAATGCGTGTGCGTGAACCGCTCCGTTGAGTCTGTCAACGCAA 3238
Db 3443 TCAACGCGCAGATCCGCGCGATCAAGGTGCGGACCGTTCGCTGAGTCCGTCACCGCCA 3502
Qy 3239 CCGCAGAAAGGAGATTCCTCCAAACAGGCGCATGTTGTCACCATTCGCTGCTGTTG 3298
Db 3503 CCGCGGAGAAAGCGATGCGCACCAAGGCGCATGTCGCGCACCATTCGCCGCTGTTG 3562
Qy 3299 TCACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAATCAATCATCG 3358
Db 3563 TCACCTGTGACCGTTCGCGAGGTGATGAGTCAAGGCTGGGACGCGGTGCCATCATTTG 3622
Qy 3359 AGGCTATGAAGATGGAAGCAACATCACTCTTCTGTTGACGGCAAGATTGAACCGCTTG 3418
Db 3623 AGGCTATGAAGATGGAAGCCACCATCACCGCGCTGTCGACGGTGTCTATCGACCGCTG 3682
Qy 3419 TGTTCTCTGCTGCACAAAGGTGGAAGTGGGACTGTCGTCGTTTCCTTA 3473
Db 3683 TGTTGCGCGCGCCACCAAGGTGAGGCGCGGACCTCATCTGTTGTTGTTCTTA 3737

RESULT 8

AAFP71419
ID AAFP71419 standard; DNA; 1719 BP.

XX AAFP71419;

XX 30-APR-2001 (first entry)

XX Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:119.
DE Corynebacterium glutamicum; carbon metabolism and energy production;
KW SMP protein; sugar metabolism and oxidative phosphorylation protein;
KW fine chemical production; organic acid; proteinogenic amino acid;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
diagnosis; Corynebacterium diphtheriae; evolutionary study; ds.

OS Corynebacterium glutamicum.

XX WO200100844-A2.

XX 04-JAN-2001.

XX 23-JUN-2000; 2000WO-IB00943.

XX 25-JUN-1999; 99US-0141031.

PR 08-JUL-1999; 99DE-1031412.

PR 08-JUL-1999; 99DE-1031413.

PR 08-JUL-1999; 99DE-1031419.

PR 08-JUL-1999; 99DE-1031420.

PR 08-JUL-1999; 99DE-1031424.

PR 08-JUL-1999; 99DE-1031428.

PR 08-JUL-1999; 99DE-1031431.

PR 08-JUL-1999; 99DE-1031433.

PR 08-JUL-1999; 99DE-1031434.

PR 08-JUL-1999; 99DE-1031510.

PR 08-JUL-1999; 99DE-1031562.

PR 08-JUL-1999; 99DE-1031634.

PR 09-JUL-1999; 99DE-1032180.
 PR 09-JUL-1999; 99DE-1032227.
 PR 09-JUL-1999; 99DE-1032310.
 PR 09-JUL-1999; 99DE-1032330.
 PR 09-JUL-1999; 99US-0433208.
 PR 14-JUL-1999; 99DE-1032524.
 PR 14-JUL-1999; 99DE-1032973.
 PR 14-JUL-1999; 99DE-1033005.
 PR 27-AUG-1999; 99DE-1040765.
 PR 31-AUG-1999; 99DE-0151572.
 PR 03-SEP-1999; 99DE-1042076.
 PR 03-SEP-1999; 99DE-1042079.
 PR 03-SEP-1999; 99DE-1042086.
 PR 03-SEP-1999; 99DE-1042087.
 PR 03-SEP-1999; 99DE-1042088.
 PR 03-SEP-1999; 99DE-1042095.
 PR 03-SEP-1999; 99DE-1042123.
 PR 03-SEP-1999; 99DE-1042125.
 XX (BADI) BASF AG.
 XX
 XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
 XX
 XX MPI: 2001-061975/07.
 DR P-PSDB; AAB79302.
 XX
 XX New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
 PT metabolism and oxidative phosphorylation protein for production or
 PT modulation of production of fine chemicals e.g. amino acids,
 PT carbohydrates or enzymes -
 XX
 XX Claim 3; Page 316-319; 1246pp; English.
 XX
 CC AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
 CC metabolism and oxidative phosphorylation (SMP) proteins given in
 CC AAB79243 to AAB 79633 which are involved in carbon metabolism and
 CC energy production. The C. glutamicum SMP gene can be used in vectors
 CC (II) for expression in host cells and production or modulation of
 CC production of fine chemicals, such as, an organic acid, a proteinogenic
 CC or nonproteinogenic amino acid (preferred), a purine or pyrimidine base,
 CC a nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty
 CC acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a
 CC cofactor, a polyketide, or an enzyme. The presence of (I) or SMP proteins
 CC (III) encoded by them are used for diagnosing the presence or activity of
 CC Corynebacterium diptheriae in a subject. (I), (II), (III) or host cells
 CC containing them are used to map genomes of organisms related to
 CC C. glutamicum, identify and localise C. glutamicum sequences of interest,
 CC in evolutionary studies, in determining SMP protein regions required
 CC for function, in modulating SMP protein activity, in modulating the
 CC metabolism of sugars, and in modulating high-energy molecule production
 CC in a cell (i.e. ATP, NADPH).
 XX
 XX Sequence 1719 BP; 405 A; 457 C; 468 G; 389 T; 0 other;
 SQ
 Query Match 45.0%; Score 1564.8; DB 22; Length 1719;
 Best Local Similarity 98.3%; Pred. No. 0;
 Matches 1592; Conservative 0; Mismatches 27; Indels 1; Gaps 1;
 QY 1 GTGACTGTATCACCCCTTGGCGGTCTCTTGTGTAAGGAATAATTACTCTAGTGTGCACT 60
 DB 101 GTGACTGTATCACCCCTTGGCGGTCTCTTGTGTAAGGAATAATTACTCTAGTGTGCACT 160
 QY 61 CACACATCTTCAACGCTTCAGACATTCAAAAGATCTTGTAGCAACCCGCGGAATC 120
 DB 161 CACACATCTTCAACGCTTCAGACATTCAAAAGATCTTGTAGCAACCCGCGGAATC 220
 QY 121 GCGGTCCGTGCTTCCGTGAGCACTCGAACCGGTGCAGCCGCTAGCTATTATACCC 180
 DB 221 GCGGTCCGTGCTTCCGTGAGCACTCGAACCGGTGCAGCCGCTAGCTATTATACCC 280
 QY 181 CGTGAAGATCGGGATCATTTCCACCGCTCTTTGCTTCTGAAGCTGTCGCAATGTACT 240
 DB 281 CGTGAAGATCGGGATCATTTCCACCGCTCTTTGCTTCTGAAGCTGTCGCAATGTACT 340

QY 241 GAAAGCTCACCAAGTCAAGCGCTTACTGTGACATCGATGAATTAATCGGTGACGTAAAAA 300
 DB 341 GAAAGCTCACCAAGTCAAGCGCTTACTGTGACATCGATGAATTAATCGGTGACGTAAAAA 400
 QY 301 GTTAAAGCAGATGCTATTATCCCGGATATGGCTTCTGCTGTAATGCAATGCCAGCTTGC 360
 DB 401 GTTAAAGCAGATGCTATTATCCCGGATATGGCTTCTGCTGTAATGCAATGCCAGCTTGC 460
 QY 361 CCGAGTCCGCGGAAAAACCGGCAATTAATTTGCGCCCAACCCGAGAGTTCTTGTATCTC 420
 DB 461 CCGAGTCCGCGGAAAAACCGGCAATTAATTTGCGCCCAACCCGAGAGTTCTTGTATCTC 520
 QY 421 ACCGGTGAATGCTCTCGTGGGTAAACCGCCGCGAAGAGGCTGCTCTCCAGTTTGGCG 480
 DB 521 ACCGGTGAATGCTCTCGTGGGTAAACCGCCGCGAAGAGGCTGCTCTCCAGTTTGGCG 580
 QY 481 GAATCCACCCCGAGCAAAAAACATCGATGACATCGTTAAAAAGCGCTGAAGGCCAGACTTAC 540
 DB 581 GAATCCACCCCGAGCAAAAAACATCGATGACATCGTTAAAAAGCGCTGAAGGCCAGACTTAC 640
 QY 541 CCCATCTTTGTAAGGCGAGTTGCGCGGTGGTGGCGGACGCGTATGCCGCTTTGTTTCTTCA 600
 DB 641 CCCATCTTTGTAAGGCGAGTTGCGCGGTGGTGGCGGACGCGTATGCCGCTTTGTTTCTTCA 700
 QY 601 CCGATGAGCTCCGCAANTTGGCAACAGACATCTCGTGAAGCTGAGCGGCAATTCGCG 660
 DB 701 CCGATGAGCTCCGCAANTTGGCAACAGACATCTCGTGAAGCTGAGCGGCAATTCGCG 760
 QY 661 GACGTTCCGTATATGTCGAACGCTGTGTATTAACCCCGACAGACATTTGAAGTGCAGATC 720
 DB 761 GATGCGCGGTATATGTCGAACGCTGTGTATTAACCCCGACAGACATTTGAAGTGCAGATC 820
 QY 721 CTTGCGCATCGCACTGGAGAAAGTTGTACACCTTTATGAACGTGACTGCTCACTCAGCGT 780
 DB 821 CTTGCGCATCGCACTGGAGAAAGTTGTACACCTTTATGAACGTGACTGCTCACTCAGCGT 880
 QY 781 CGTCAACAAAAAGTTGTGCAAAATTCGCCAGCACAGCAATTTGGATCCAGAACTCGCGTAT 840
 DB 881 CGTCAACAAAAAGTTGTGCAAAATTCGCCAGCACAGCAATTTGGATCCAGAACTCGCGTAT 940
 QY 841 CGCATTTCTGCGGATGACAGTAAAGTTCTGCCGCTCCATTTGTTTACCAGGCGCGGGAACC 900
 DB 941 CGCATTTCTGCGGATGACAGTAAAGTTCTGCCGCTCCATTTGTTTACCAGGCGCGGGAACC 1000
 QY 901 GTGGAATTTCTGGTGCATGAAAGGCAACACCGTTTTCATCGAAATGAACCCAGCTATC 960
 DB 1001 GTGGAATTTCTGGTGCATGAAAGGCAACACCGTTTTCATCGAAATGAACCCAGCTATC 1060
 QY 961 CAGGTTGAGCAACCGTGCACTGAGAAAGTCAACGAGGTGAGCTGCTGTAAGGCGCAGATG 1020
 DB 1061 CAGGTTGAGCAACCGTGCACTGAGAAAGTCAACGAGGTGAGCTGCTGTAAGGCGCAGATG 1120
 QY 1021 CGTTGGCTGTGGTGCACCTTGAAGGAATTTGGGTCTGACCCCAAGATAAGATCAAGACC 1080
 DB 1121 CGCTTGGCTGTGGTGCACCTTGAAGGAATTTGGGTCTGACCCCAAGATAAGATCAAGACC 1180
 QY 1081 CACGTTGAGCACTGCACTGCGCATCAACGGAAGATCCAAACCAACCGGTTCCGCCCA 1140
 DB 1181 CACGTTGAGCACTGCACTGCGCATCAACGGAAGATCCAAACCAACCGGTTCCGCCCA 1240
 QY 1141 GATACCGAACTATCAACCGGTACCGCTCACAGCGAGCTGGCGTTCTGTTGACCGT 1200
 DB 1241 GATACCGAACTATCAACCGGTACCGCTCACAGCGAGCTGGCGTTCTGTTGACCGT 1300
 QY 1201 GCAGCTCAGCTCGGTGGGAAATCACCGCACACTTTGATCCATGCTGTGTAATGAC 1260
 DB 1301 GCAGCTCAGCTCGGTGGGAAATCACCGCACACTTTGATCCATGCTGTGTAATGAC 1360
 QY 1261 TCCGTTGGTTCGACTTTGAAACTGCTGTTGTCGTCAGACGCGGCTTGGCTGAGTTC 1320
 DB 1361 TCCGTTGGTTCGACTTTGAAACTGCTGTTGTCGTCAGACGCGGCTTGGCTGAGTTC 1420
 QY 1321 ACCGTTGCTGGTGTGGCAACCAACATTTGCTTCTTGGCTGCGTTCGCTCGGGAAGAGAC 1380

PR	03-SEP-1999;	99DE-1042095.	
PR	03-SEP-1999;	99DE-1042123.	
PR	03-SEP-1999;	99DE-1042125.	
XX			
XX	(BADI) BASF AG.		
PA			
XX	Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;		
PI	WPI; 2001-061975/07.		
XX	P-PSDB; AAB79303.		
DR			
DR			
XX			
XX	New isolated Corynebacterium glutamicum nucleic acid encoding a sugar		
PT	metabolism and oxidative phosphorylation protein for production or		
PT	modulation of production of fine chemicals e.g. amino acids,		
PT	carbohydrates or enzymes -		
XX			
PS	Claim 3; Page 321-323; 1246pp; English.		
XX			
CC	AAAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar		
CC	metabolism and oxidative phosphorylation (SMP) proteins given in		
CC	AAB79243 to AAB 79633 which are involved in carbon metabolism and		
CC	energy production. The C. glutamicum SMP gene can be used in vectors		
CC	(II) for expression in host cells and production or modulation of		
CC	production of fine chemicals, such as, an organic acid, a proteinogenic		
CC	or nonproteinogenic amino acid (preferred), a purine or pyrimidine base,		
CC	a nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty		
CC	acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a		
CC	cofactor, a polyketide, or an enzyme. The presence of (I) or SMP proteins		
CC	(III) encoded by them are used for diagnosing the presence or activity of		
CC	Corynebacterium diptheriae in a subject. (II), (III) or host cells		
CC	containing them are used to map genomes of organisms related to		
CC	C. glutamicum, identify and localise C. glutamicum sequences of interest,		
CC	in evolutionary studies, in determining SMP protein regions required		
CC	for function, in modulating SMP protein activity, in modulating the		
CC	metabolism of sugars, and in modulating high-energy molecule production		
CC	in a cell (i.e. ATP, NADPH).		
XX			
XX	Sequence 1406 BP; 325 A; 381 C; 385 G; 315 T; 0 other;		
SQL			
	Query Match	38.9%; Score 1351.8; DB 22; Length 1406;	
	Best Local Similarity	98.0%; Pred. No. 0;	
	Matches 1379; Conservative	0; Mismatches 27; Indels 1; Gaps 1;	
QY	214	GCTTCTGAAGCTGTCCGATGGTACTCAAGGCTCACAGTCAAGGCGTACTCGACATC	273
DB	1	GCTTCTGAAGCTGTCCGATGGTACTCAAGGCTCACAGTCAAGGCGTACTCGACATC	60
QY	274	GATGAATTTATCGGTGCAGCTAAAGAGTTAAAGCAGATGCTATTTACCGGGATATGGC	333
DB	61	GATGAATTTATCGGTGCAGCTAAAGAGTTAAAGCAGATGCCATTTACCGGGATACGGC	120
QY	334	TTCTGTCTGAAATATGCCAGTGTGCCCGCAGTGCAGGAAACGGCATTACTTTTATT	393
DB	121	TTCTGTCTGAAATATGCCAGTGTGCCCGCAGTGTCCGAAACGGCATTACTTTTATT	180
QY	394	GGCCCAACCCAGAGGTTCTTGATCTCACCGTGTAAAGTCTGTCGGGTAAACCGCCGC	453
DB	181	GGCCCAACCCAGAGGTTCTTGATCTCACCGTGTAAAGTCTGTCGGGTAAACCGCCGC	240
QY	454	AAGAAGGCTGGTCTGCCAGTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGATC	513
DB	241	AAGAAGGCTGGTCTGCCAGTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGATC	300
QY	514	GTTTAAAGCGCTGAAGGCAGACTTACCCCATCTTTGTAAAGCAGTTGCCGTTGTTGGC	573
DB	301	GTTTAAAGCGCTGAAGGCAGACTTACCCCATCTTTGTAAAGCAGTTGCCGTTGTTGGC	360
QY	574	GGACCGGTTATCGGCTTTTGTCTTTCACCTCATGAGTCCGCAATTTGGCAACAGAGCA	633
DB	361	GGACCGGTTATCGGCTTTTGTCTTTCACCTCATGAGTTCGCAATTTAGCAACAGAGCA	420
QY	634	TCTCGTGAAGCTGAAGCGCATTTCCGGCAGCGTTTCGGTATATGTCGAACGTTGTGATT	693

CC acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a
CC cofactor, a polyketide, or an enzyme. the presence of (i) or SMP proteins
CC (iii) encoded by them are used for diagnosing the presence or activity of
CC Corynebacterium diptheriae in a subject. (ii), (iii) or host cells
CC containing them are used to map genomes of organisms related to
CC C. glutamicum, identify and localise C. glutamicum sequences of interest,
CC in evolutionary studies, in determining SMP protein regions required
CC for function, in modulating SMP protein activity, in modulating the
CC metabolism of sugars, and in modulating high-energy molecule production
CC in a cell (i.e. ATP, NADPH).
XX
SQ Sequence 1083 BP; 211 A; 343 C; 300 G; 229 T; 0 other;

Query Match 29.4%; Score 1020.2; DB 22; Length 1083;
Best Local Similarity 98.4%; Pred. No. 1.2e-268;
Matches 1062; Conservative 0; Mismatches 13; Indels 4; Gaps 3;

QY 1597 CTGAAGCAGCTTGCCCGCAGCGCTTCTCGTGTATCTCCGTGAGCAGCGCAGCTGGCA 1656
DB 5 CTGAAGCAGCTTGCCCGCAGCGCTTCTCGTGTATCTCCGTGAGCAGCGCAGCTGGCA 64
QY 1657 GTTACTGATACCACTTCCCGGATGCACACCAAGTCTTTGTGTCGACCCGAGTCCGCTCA 1716
DB 65 GTTACTGATACCACTTCCCGGATGCACACCAAGTCTTTGTGTCGACCCGAGTCCGCTCA 124
QY 1717 TTCGACTGAAGCTGGCGGAGAGCGCTGCGCAAGCTGACTCTGAGCTTTTGTCCGTG 1776
DB 125 TTCGACTGAAGCTGGCGGAGAGCGCTGCGCAAGCTGACTCTGAGCTTTTGTCCGTG 184
QY 1777 GAGCCTGGGCGCGGACCTAGCATGTGGGATGGTTCCTCTTTGAGGATCCGTGG 1836
DB 185 GAGCCTGGGCGCGGACCTAGCATGTGGGATGGTTCCTCTTTGAGGATCCGTGG 244
QY 1837 GACAGGCTCGAGAGCTGCGGAGCGGATGCCGAATGTAAACATTTCAGATCTCTCTCGC 1896
DB 245 GACAGGCTCGAGAGCTGCGGAGCGGATGCCGAATGTAAACATTTCAGATCTCTCTCGC 304
QY 1897 GCGCGCAACCGTGGGATATACCCCGGTACCCAGACTCCCGTCTGCGCGCGCTTTGTTAAG 1956
DB 305 GCGCGCAACCGTGGGATATACCCCGGTACCCAGACTCCCGTCTGCGCGCGCTTTGTTAAG 364
QY 1957 GAAGCTGCCAGCTCCGCGGTGGATCTTCCGATCTTCGAGCGCTTAAAGAGCTCTCC 2016
DB 365 GAAGCTGCCAGCTCCGCGGTGGATCTTCCGATCTTCGAGCGCTTAAAGAGCTCTCC 424
QY 2017 CAGATGGCTCCAGCAATCGACGAGTCTCTGAGACCAACACCGCGGTAGCGAGTGGCT 2076
DB 425 CAGATGGCTCCAGCAATCGACGAGTCTCTGAGACCAACACCGCGGTAGCGAGTGGCT 484
QY 2077 ATGGCTTATCTGTGTATCTCTGTATCCAAATGAAAGCTCTACACCTGGATTACTAC 2136
DB 485 ATGGCTTATCTGTGTATCTCTGTATCCAAATGAAAGCTCTACACCTGGATTACTAC 544
QY 2137 CTAAAGATGGCAGAGGATCGTCAAGTCTGGCGCTACATCTGGCCATTAAAGATATG 2196
DB 545 CTAAAGATGGCAGAGGATCGTCAAGTCTGGCGCTACATCTGGCCATTAAAGATATG 604
QY 2197 GCTGGTCTGCTTCCGCCAGCTGCGGTAAACAAAGTGTACCGCACTGCGCGGTGAATTC 2256
DB 605 GCTGGTCTGCTTCCGCCAGCTGCGGTAAACAAAGTGTACCGCACTGCGCGGTGAATTC 664
QY 2257 GATCTGCAGTGCACTGTCACACCCAGCACTCGGGTGGCAGTGTGGCTACCTACTTT 2316
DB 665 GATCTGCAGTGCACTGTCACACCCAGCACTCGGGTGGCAGTGTGGCTACCTACTTT 724
QY 2317 GCTGCAGTCAAGTGGTGCAGATGCTGTGACCGTGTCTCC-CCACCACTGTCTGGCAC 2375
DB 725 GCTGCAGTCAAGTGGTGCAGATGCTGTGACCGTGTCTCCCGGCACTGTCTGGCAC 784
QY 2376 CACCTCCCA--GCCATCCCTGTCTGCCATTGTTCTGATTCGCGCACACCCGTCGCGAT 2433
DB 785 CACCTCCCAAGCCATCCCTGTCTGCCATTGTTCTGATTCGCGCACACCCGTCGCGAT 844

QY 2434 ACCGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCTGTGCGCGGA 2493
DB 845 ACCGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGTGCGCGGA 904
QY 2494 CTGTACCTGCGCATTTGAGTCTGGAAACCCAGGCCAACCCGTCGCTCTACCGCCACGAA 2553
DB 905 CTGTACCTGCGCATTTGAGTCTGGAAACCCAGGCCAACCCGTCGCTCTACCGCCACGAA 964
QY 2554 ATCCAGGCGGACAGTTGTCTCAACCTCGCTGTCACAGCCACCGCACTTGGGCTTGTGAT 2613
DB 965 ATCCAGGCGGACAGTTGTCTCAACCTCGCTGTCACAGCCACCGCACTTGGGCTTGTGAT 1024
QY 2614 CGCTTCGAGCTCATCGAAGACACTACGC-AGCGTTAATGAGATGCTGGGACGCCCAA 2671
DB 1025 CGTTTCGAACCTCATCGAAGACAACTACGCAAGCGGTTAATGAGATGCTGGGACGCCCAA 1083
RESULT 11
AAF71418
ID AAF71418 standard; DNA; 1083 BP.
XX AC AAF71418;
XX AC AAF71418;
DT 30-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:117.
KW Corynebacterium glutamicum; carbon metabolism and energy production;
KW SMP protein; sugar metabolism and oxidative phosphorylation protein;
KW fine chemical production; organic acid; proteinogenic amino acid;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KW diagnosis; Corynebacterium diptheriae; evolutionary study; ds.
XX
OS Corynebacterium glutamicum.
XX
PN WO200100844-A2.
XX 04-JAN-2001.
XX 23-JUN-2000; 2000WO-IB00943.
XX 25-JUN-1999; 99US-0141031.
PR 08-JUL-1999; 99DE-1031412.
PR 08-JUL-1999; 99DE-1031413.
PR 08-JUL-1999; 99DE-1031419.
PR 08-JUL-1999; 99DE-1031420.
PR 08-JUL-1999; 99DE-1031424.
PR 08-JUL-1999; 99DE-1031428.
PR 08-JUL-1999; 99DE-1031431.
PR 08-JUL-1999; 99DE-1031433.
PR 08-JUL-1999; 99DE-1031434.
PR 08-JUL-1999; 99DE-1031510.
PR 08-JUL-1999; 99DE-1031562.
PR 08-JUL-1999; 99DE-1031634.
PR 09-JUL-1999; 99DE-1032180.
PR 09-JUL-1999; 99DE-1032227.
PR 09-JUL-1999; 99DE-1032230.
PR 09-JUL-1999; 99US-0143208.
PR 14-JUL-1999; 99DE-1032924.
PR 14-JUL-1999; 99DE-1032973.
PR 14-JUL-1999; 99DE-1033005.
PR 27-AUG-1999; 99DE-1040765.
PR 31-AUG-1999; 99US-0151572.
PR 03-SEP-1999; 99DE-1042076.
PR 03-SEP-1999; 99DE-1042079.
PR 03-SEP-1999; 99DE-1042086.
PR 03-SEP-1999; 99DE-1042087.
PR 03-SEP-1999; 99DE-1042088.
PR 03-SEP-1999; 99DE-1042095.
PR 03-SEP-1999; 99DE-1042123.
PR 03-SEP-1999; 99DE-1042125.

Db	121	GGTCAACCCCATCTCTCCAAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGT	180
Qy	2736	AGATCCAGCAGACTTTGCTGCAGACCCACAAAAGTAGACATCCCGAGACTCTGTCTCATCGC	2795
Db	181	GGATCCAGCAGACTTTGCTGCCGATCCAAAGTAGACATCCCGAGACTCTGTCTCATCGC	240
Qy	2796	GTTCTCTGGCGGAGCTTGGTAACCTCCAGGTGGCTGGCGAGAACCACTCGCACCG	2855
Db	241	GTTCTCTGGCGGAGCTTGGTAACCTCCAGGTGGCTGGCGAGACCACTCGCACCG	300
Qy	2856	CGCACTGGAGGCGCTCCGAAAGGCAAGCACCCTCTGACGGAAGTTCTCTGAGGAAGAGCA	2915
Db	301	CGCACTGGAGGCGCTCCGAAAGGCAAGCACCCTCTGACGGAAGTTCTCTGAGGAAGAGCA	360
Qy	2916	GGCGACCTCGACGCTGATGATTCCAAAGGAAAGTCGCAACAGCCTCAACCGCTCTCTGTT	2975
Db	361	GGCGACCTCGACGCTGATGATTCCAAAGGAAAGTCGCAATAGCCTCAACCGCTCTCTGTT	420
Qy	2976	CCCGAAGCCAAACGAAGAGTTCTCGAGCAACGTCGCGCTTTGCGCAACACCTCTGCGCT	3035
Db	421	CCCGAAGCCAAACGAAGAGTTCTCGAGCAACGTCGCGCTTTGCGCAACACCTCTGCGCT	480
Qy	3036	GGATGATCGTGAATTTCTTACCGACTGTCGAGGGCCGAGACTTTGATCCGCGCTGCC	3095
Db	481	GGATGATCGTGAATTTCTTACCGCGCTGGTCGAAGGCGCGAGACTTTGATCCGCGCTGCC	540
Qy	3096	AGATGTGCGACCCCACTGCTTCTGCGCTGGATGCGATCTCTGAGCCAGACGATAAGG	3155
Db	541	AGATGTGCGACCCCACTGCTTCTGCGCTGGATGCGATCTCTGAGCCAGACGATAAGG	600
Qy	3156	TATCGCAATTTGTTGGCCAAACGTCAACGGCCAGATCCGCCCAATCGCTGTGCGTGACCG	3215
Db	601	TATCGCAATTTGTTGGCCAAACGTCAACGGCCAGATCCGCCCAATCGCTGTGCGTGACCG	660
Qy	3216	CTCGGTTGAGTCTGTACCGCAACCCGAGAAAGGAGATTCTCTCAACAGGGGCAATGT	3275
Db	661	CTCGGTTGAGTCTGTACCGCAACCCGAGAAAGGAGATTCTCTCAACAGGGGCAATGT	720
Qy	3276	TGCTGCACCAATTCGCTGGTGTCTGCTACTGTGCTGCTGAAAGTGATGAGGTCAAGGC	3335
Db	721	TGCTGCACCAATTCGCTGGTGTCTGCTACTGTGCTGCTGAAAGTGATGAGGTCAAGGC	780
Qy	3336	TGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCCTGCTTCTGT	3395
Db	781	TGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCCTGCTTCTGT	840
Qy	3396	TGACGGCAAGATTGAAGCGGTTGTGTTCTGTCGAACGAAGGTGGAAGTGGCGACTT	3455
Db	841	TGACGGCAAAATCGATCGCGTTGTGTTCTGTCGAACGAAGGTGGAAGTGGCGACTT	900
Qy	3456	GATCGTCGCTGTTTCTCTAA	3474
Db	901	GATCGTCGCTGTTTCTCTAA	919
RESULT 13			
ID AAF71416			
ID AAF71416 standard; DNA; 939 BP.			
XX	AAF71416;		
XX	AC		
XX	DT		
XX	30-APR-2001 (first entry)		
DE	Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:113		
KW	Corynebacterium glutamicum; carbon metabolism and energy production;		
KW	SMP protein; sugar metabolism and oxidative phosphorylation protein;		
KW	fine chemical production; organic acid; proteinogenic amino acid;		
KW	nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;		
KW	nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;		
KW	carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;		
XX	diagnosis; Corynebacterium diphtheriae; evolutionary study; ds.		

OS Corynebacterium glutamicum.

XX PN WO200100844-A2.

XX PD 04-JAN-2001.

XX PF 23-JUN-2000; 2000WO-IB00943.

XX PR 25-JUN-1999; 99US-0141031.

XX PR 08-JUL-1999; 99DE-1031412.

XX PR 08-JUL-1999; 99DE-1031413.

XX PR 08-JUL-1999; 99DE-1031419.

XX PR 08-JUL-1999; 99DE-1031420.

XX PR 08-JUL-1999; 99DE-1031424.

XX PR 08-JUL-1999; 99DE-1031428.

XX PR 08-JUL-1999; 99DE-1031431.

XX PR 08-JUL-1999; 99DE-1031433.

XX PR 08-JUL-1999; 99DE-1031434.

XX PR 08-JUL-1999; 99DE-1031510.

XX PR 08-JUL-1999; 99DE-1031562.

XX PR 08-JUL-1999; 99DE-1031634.

XX PR 09-JUL-1999; 99DE-1032180.

XX PR 09-JUL-1999; 99DE-1032227.

XX PR 09-JUL-1999; 99DE-1032230.

XX PR 09-JUL-1999; 99US-0143208.

XX PR 14-JUL-1999; 99DE-1032973.

XX PR 14-JUL-1999; 99DE-1033005.

XX PR 27-AUG-1999; 99DE-1040765.

XX PR 31-AUG-1999; 99US-0515172.

XX PR 03-SEP-1999; 99DE-1042076.

XX PR 03-SEP-1999; 99DE-1042079.

XX PR 03-SEP-1999; 99DE-1042086.

XX PR 03-SEP-1999; 99DE-1042087.

XX PR 03-SEP-1999; 99DE-1042088.

XX PR 03-SEP-1999; 99DE-1042095.

XX PR 03-SEP-1999; 99DE-1042123.

XX PR 03-SEP-1999; 99DE-1042125.

XX PA (BADI) BASF AG.

XX PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauser G;

XX XX WPI; 2001-061975/07.

XX DR P-PSDB; AAB79299.

XX New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
 PT metabolism and oxidative phosphorylation protein for production or
 PT modulation of production of fine chemicals e.g. amino acids,
 PT carbohydrates or enzymes -

XX Claim 3; Page 308-310; 1246pp; English.

XX AAB71360 to AAB71750 encode the Corynebacterium glutamicum sugar
 CC metabolism and oxidative phosphorylation (SMP) proteins given in
 CC AAB79243 to AAB 79633 which are involved in carbon metabolism and
 CC energy production. The C. glutamicum SMP gene can be used in vectors
 CC (II) for expression in host cells and production or modulation of
 CC production of fine chemicals, such as, an organic acid, a proteinogenic
 CC or nonproteinogenic amino acid (preferred), a purine or pyrimidine base,
 CC a nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty
 CC acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a
 CC cofactor, a polyketide, or an enzyme. The presence of (I) or SMP proteins
 CC (III) encoded by them are used for diagnosing the presence or activity of
 CC Corynebacterium diptheriae in a subject. (I), (II), (III) or host cells
 CC containing them are used to map genomes of organisms related to
 CC C. glutamicum, identify and localise C. glutamicum sequences of interest,
 CC in evolutionary studies, in determining SMP protein regions required
 CC for function, in modulating SMP protein activity, in modulating the
 CC metabolism of sugars, and in modulating high-energy molecule production
 CC in a cell (i.e. ATP, NADPH).

XX Sequence 939 BP; 202 A; 283 C; 261 G; 193 T; 0 other;

Query Match	25.4%;	Score 883;	DB 22;	Length 939;
Best Local Similarity	98.3%;	Pred. No. 3.9e-231;		
Matches	903;	Conservative	Mismatches 15;	Indels 1; Gaps 1;
QY	2557	CCAGGCGGACAGTTGTCCAACTCGGTGCAACAGGCGACCGACCTGGGCGCTTGTGATCGC	2616	
DB	1	CCAGGCGGACAGTTGTCCAACTCGGTGCAACAGGCGACCGACCTGGGCGCTTGTGATCGT	60	
QY	2617	TTGAGCTCATCGAAGCAACTACGC-AGCGTTAATGAGATGCTGGAGCGCCCAACAA	2675	
DB	61	TTGAACTCATCGAAGCAACTACGC-AGCGTTAATGAGATGCTGGAGCGCCCAACAA	120	
QY	2676	GGTACCCCATCTCTCAAGGTTGTGGGACCTCGCACTCCACCTGTTGGTGGGGTGT	2735	
DB	121	GGTACCCCATCTCTCAAGGTTGTGGGACCTCGCACTCCACCTGTTGGTGGGGTGT	180	
QY	2736	AGATCCAGCAGACTTGTGTCAGACCCCAAAAGTAGACATCCACAGACTTGTTCATCGC	2795	
DB	181	GGATCCAGCAGACTTGTGTCGCGATCCAAAGTAGACATCCACAGACTTGTTCATCGC	240	
QY	2796	GTTCCTGCGCGCGAGCTTGGTAAACCTTCCAGTGGTGGCGACAAACACTGCGACCCG	2855	
DB	241	GTTCCTGCGCGCGAGCTTGGTAAACCTTCCAGTGGTGGCGACAGCCACTGCGACCCG	300	
QY	2856	CGCACTGGAAGCGCGCTCCGGAAGCAAGGCACCTCTGACGGAAGTTCTTGAGGAAGACA	2915	
DB	301	CGCACTGGAAGCGCGCTCCGGAAGCAAGGCACCTCTGACGGAAGTTCTTGAGGAAGACA	360	
QY	2916	GGCGCACCTCGACGCTGATGATTCGAAGGAACGTCGCAACAGCCTCAACCGCTGCTGTT	2975	
DB	361	GGCGCACCTCGACGCTGATGATTCGAAGGAACGTCGCAACAGCCTCAACCGCTGCTGTT	420	
QY	2976	CCGGAAGCAACCGAAGAGTTCTTCGAGCACCGTCCGCGCTTCGCGCAACACTCTGCGCT	3035	
DB	421	CCGGAAGCAACCGAAGAGTTCTTCGAGCACCGTCCGCGCTTCGCGCAACACTCTGCGCT	480	
QY	3036	GGATGATGTAATTTCTTACGACTGCTGAGGCGCGCGAGACTTTGATCCGCTGCC	3095	
DB	481	GGATGATGTAATTTCTTACGACTGCTGAGGCGCGCGAGACTTTGATCCGCTGCC	540	
QY	3096	AGATGCGGACCCCACTGCTGTTGCTGGCTGATGCGATCTCTGAGCCAGACGATAGGG	3155	
DB	541	AGATGCGGACCCCACTGCTGTTGCTGGCTGATGCGATCTCTGAGCCAGACGATAGGG	600	
QY	3156	TATGCGCAATGTTGGCCAAAGTCACGCGCAGATCCGCCCAATGCTGCGTGACCG	3215	
DB	601	TATGCGCAATGTTGGCCAAAGTCACGCGCAGATCCGCCCAATGCTGCGTGACCG	660	
QY	3216	CTCCGCTTGAGTCTGTACCGCAACCGCAGAAAAGGACAGATTCTCTCCAAAGGGCCATGT	3275	
DB	661	CTCCGCTTGAGTCTGTACCGCAACCGCAGAAAAGGACAGATTCTCTCCAAAGGGCCATGT	720	
QY	3276	TGCTGCACCATTCGCTGTTGTTGTCACGTGACTGTTGCTGAAAGGTGATGAGGTCAAGGC	3335	
DB	721	TGCTGCACCATTCGCTGTTGTTGTCACGTGACTGTTGCTGAAAGGTGATGAGGTCAAGGC	780	
QY	3336	TGAGATCGAGTCGCAATCATCGAGGCTATGAAGATGAAGCAACAACTCACTGCTTCTGT	3395	
DB	781	TGAGATCGAGTCGCAATCATCGAGGCTATGAAGATGAAGCAACAACTCACTGCTTCTGT	840	
QY	3396	TCAGGCGAAGATTGAACGGTGTGGTTCTGCTGCAAGAGGTGGAAGTGGCGACTT	3455	
DB	841	TCAGGCGAAGATTGAACGGTGTGGTTCTGCTGCAAGAGGTGGAAGTGGCGACTT	900	
QY	3456	GATGCTGCTGCTTCTCTAA	3474	
DB	901	GATGCTGCTGCTTCTCTAA	919	
RESULT 14				
ABK63659				
ID	ABK63659	standard; cDNA; 3945 BP.		

XX ABK63659;
 XX 18-JUN-2002 (first entry)
 XX Rat sequence differentially expressed in response to a hepatotoxin #1566.
 XX Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
 XX differential expression; centrilobular necrosis; steatosis.
 XX Rattus norvegicus.
 XX WO200210453-A2.
 XX 07-FEB-2002.
 XX 30-JUL-2001; 2001WO-US23872.
 XX 31-JUL-2000; 2000US-222040P.
 XX 02-NOV-2000; 2000US-244880P.
 XX 11-MAY-2001; 2001US-290029P.
 XX 15-MAY-2001; 2001US-290645P.
 XX 22-MAY-2001; 2001US-292336P.
 XX 06-JUN-2001; 2001US-295798P.
 XX 13-JUN-2001; 2001US-297457P.
 XX 19-JUN-2001; 2001US-298884P.
 XX 09-JUL-2001; 2001US-303459P.
 XX (GENE-) GENE LOGIC INC.
 XX Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;
 XX WPI; 2002-241625/29.
 XX Predicting toxic effects of compounds or the progression of these toxic
 XX effects by determining the changes in gene expression in tissues or
 XX cells exposed to the toxin and comparing these to gene expression in
 XX unexposed tissues or cells -
 XX Claim 1; Seq ID No 1566; 239pp; English.
 XX The invention relates to methods for predicting toxic effects of
 XX compounds or the progression of these toxic effects by determining the
 XX global changes in gene expression in tissues or cells exposed to the
 XX toxin and comparing these to gene expression in unexposed tissues or
 XX cells. Also included are methods of predicting at least one toxic
 XX effect of a compound or progression of a toxic effect, preferably the
 XX hepatotoxicity of a compound, comprising detecting the level of
 XX expression in a tissue or cell sample exposed to the compound of two or
 XX more genes listed in the specification, where differential expression of
 XX the genes is indicative of at least one toxic effect or progression.
 XX The method can also be used to identify an agent which modulates the
 XX toxic response and predict cellular pathways that a compound modulates
 XX in a cell. The methods utilise a set of at least two probes (on a solid
 XX support in kit form), where each of the probes comprises a sequence that
 XX specifically hybridises to a gene listed in the specification, a computer
 XX system comprising a database containing information identifying the
 XX expression level in a tissue or cell sample exposed to a hepatotoxin of a
 XX set of genes comprising at least two genes listed in the specification,
 XX and a user interface to view the information used to present information
 XX identifying the expression level in a tissue or cell of at least one gene
 XX listed in the specification. The method is useful for elucidating global
 XX changes in gene expression and for identifying toxicity markers in
 XX tissues or cell exposed to a known toxin. The genes may be used as
 XX toxicity markers in drug screening and toxicity assays. The genes and
 XX gene expression information may be used as diagnostic markers for the
 XX prediction or identification of the physiological state of tissue or cell
 XX sample that has been exposed to a compound or agent. Hepatotoxicity
 XX is characterised by centrilobular necrosis and steatosis. The present
 XX sequence is an expressed sequence tag (EST) or cDNA derived from a gene
 XX which is differentially expressed in response to a hepatotoxic agent.

XX Sequence 3945 BP; 883 A; 1108 C; 1127 G; 827 T; 0 other;
 XX SQ

Query Match 19.6%; Score 681; DB 24; Length 3945;
 Best Local Similarity 54.2%; Pred. NO. 1.6e-175;
 Matches 1521; Conservative 0; Mismatches 1260; Indels 24; Gaps 6;
 QY 86 TCATAAAGATCTTGGTAGCAACCGCGCGAATCGCGTCCGCTTCGGTGCAGCAC 145
 DB 144 TCAGAAGTAAGTGGTGCCCAACAGAGGTGAGATTGCCATCCGAGTGTTCGTGCCCTGCA 203
 QY 146 TCGAAACCGGTGCAGCCACCGTAGCTATTTACCCCGCTGAAGATCGGGATCATTTCCACC 205
 DB 204 CAGAGCTGGGTATCCGCACAGTGGCTGTCTACTCGGAGCAGACACAGCCAGATGCACC 263
 QY 206 GCTCTTTTGTCTTGAAGCTGTCCGCAATGGTACTGAAGGCTCACCAAGTCAAGCGGTACC 265
 DB 264 GGCAGAAAGCTGATGAAGCCTACCTTATTGGCCGTGGCTGCTCTGTGCAAGCCTACC 323
 QY 266 TGGACATCGATGAATATTCGGTGCAGCTAAAGATTAAGAGCAGATGCTATTTACCGG 325
 DB 324 TGCACATTCAGACATCATTAAGGTGGCCAAAGGAGAAATGGTGTAGATGCTGTGCACCTG 383
 QY 326 GATATGGCTTCCTGTCTGAAAATGCCAGCTTGCCCGGAGTGGCGGAAAAACGGCATT 385
 DB 384 GCTATGGGTTCTCTCAGAGAGAGCAGACTTTGCCAGGCTGCCAAGATGCTGGAGTCC 443
 QY 386 CTTTATTGGCCCAACCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCTGTCGGTAA 445
 DB 444 GATTTCATTGGTCCCAAGCCAGAGGTGGTCCGCAAGATGGGAGACAAGGTGGAAGCCCGGG 503
 QY 446 CGCCCGGAAGAAGGCTGGTCTGCCAGTTTGCCCGGAGTGGCGGAAAAACGGCATT 502
 DB 504 CCATTGCCATTGTCTGAGGCGTTCCAGTGGTCTCTGCACTAATTTCCCGCATCAATTTCCC 563
 QY 503 TCATATGACATCGTTAAAGCGCTGAAGGCCAGACTTACCCCATCTTTTGAAGGACGTTG 562
 DB 564 TGATAGGGCAGCAGGAGTTCTTAACACCTATGGTTTCCCTATTATCTTCAAGGCTGCCT 623
 QY 563 CCGGTGGTGGCGGACGCGGTATGCGTTTGTCTTCCCTCATGATAGCTCCGCAATTTGG 622
 DB 624 ATGAGGTGGGGCGGTGGCATGAGGGTTGTGCATAGTAGTACGAGGAGCTGGAAGAGATT 683
 QY 623 CAACAGNAGCATCTCGTGAAGCTGAAGCGCATTCGGGACGTTTCGGTATATGTGCAAC 682
 DB 684 ACACCCGGGCTTACCTCGAGGCTTGGCAGCCTTTGGGAATGGGCAATTTGTTGTGAGA 743
 QY 683 GTGCTGTGATTAAACCCCAACACCATTTGAAGTGAGATCTTTGGCGATCGCACTGGAGAAG 742
 DB 744 AATTCAITGAGAAGCCAAAGACACATTGAGGTGCAGATCTCTAGGGACCAATATGGGAACA 803
 QY 743 TTGTACACCTTTTGAACGTTGACTGCTCACTGCAGCGTCTGCACCAAAAAGTTGTGAAA 802
 DB 804 TCTTGCACTTGTATGAGCGGGAAGTGTCCATCCAGCGCGGCGCACAGAGGTGTGAGAGA 863
 QY 803 TTGGCCAGACAGCATTTGGATTCAGAACTGCGTATGCCATTTGTGCGGATGCAAGTAA 862
 DB 864 TTGCCCTGTACCCACTGGACCCCACTTCGGTCAAGCTCACCGCTCACTCTGTCA 923
 QY 863 AGTTCTGCCGCTCCATTGGTTTACCAGGGCGGGAACCGTGGAAATTTCTTGGTGCATGAAA 922
 DB 924 AATTGCCAAGCAGGTGGCTATGAGAAATGAGGCACTGGGAGTTCTCTGGTGACAAGC 983
 QY 923 AGGGCAACCACTGTTTTCATCGAAATGAACCCAGCTATCCAGTTTGAGCACACCGTGAAGT 982
 DB 984 ATGCAAGCAGCTACTTCAATCGAGGTCAATTTCCCGCTGAGGTGGAGCACACCGTCACTG 1043
 QY 983 AAGAAGTACCGAGGTGGACCTCGTGAAGGCGCAGATGCGCTTTGGTGTGGTGCAACCT 1042
 DB 1044 AGGAGATTACAGATGTGGACCTGGTCCATGCTCAGATCCATGTGTCCGAAGGCGGAGCC 1103
 QY 1043 TGAAGNAATTGGTCTTGACCAAGATAAGATCAAGACCCAGCGTGCAGCACTGCAGTGCC 1102
 DB 1104 TGCTGACCTAGGCGCTGCGGCAAGNAACATCCGAATCAATGGTTGTGTCATTGAGTGTC 1163

1103 GCATCACCGAGATCCAAACAACGGCTTCGCCCCAGATACCGGAACATATCACCGCT 1162
1164 GGGTCACCATGAGACCCCTGCAGAGCTTCACGACAGACATGCGCGCATGAGGTTT 1223
1163 ACCGCTCACACGGGAGCTGGCGTTGCTTGAAGG---CAGCTCAGCTCGGTGGG 1219
1224 TCCGAGTGGTGAAGGCATGAGGCATCCGCTGGACAATGCTCAGCAATTCAGGGAGCTG 1283
1220 AAATCACCGCACATTTGACTCCATGCTGGTGAATAATGAACTGCGCGGTGCTCCGATTTG 1279
1284 TCATATCCCCCACTATGACTCCCTGCTCGTCAAAAGTCATTGCCCCATGGCAAGACACC 1343
1280 AAATGCTGTTGCTGTCGACAGCGCGCTTGGCTGAGTTTCAACGCTGCTGGTGTGCA 1339
1344 CTACAGCTGCCACCAAGATGAGCAGAGCCCTTGGCGAGTTCCGCTGCCGAGGTAAAGA 1403
1340 CCAACATTTGGTTTCTTGGCTGCTTGGGGGAGAGGACTTCACTTCCAAAGCGCATCG 1399
1404 CCAACATCCCTTCTGCAGAAATGCTCAACAACAGCAGTTCCTAGCGGGCATTTGG 1463
1400 CCACCGGATTTATCGCGGATCACCCACACCTCTTCAAGGCTCCACCTCGCGGATGATGAGC 1459
1464 ACACCCAGTTTCATCGATGAGAACCCCGAGCTGTTCCAGCTCGCGCTGCACAGAACCGGG 1523
1460 AGGACGATCTTGATTAAGTTCAGATGTCACCGTGAACAGCCTCATGCTGCGTC 1519
1524 CCGAAGTGTCTACATTAACCTTGGACAGCTGATGCTGATGCTGCTGCTGCTGCTGCTGCT 1583
1520 CAAAGGATTTGACGACCAACATGATGAAGC-----TGCCCAACATCAAGGATCTGCCAC 1573
1584 CCGTCAAGGTCAGTCCAGACCTTGGACCCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1643
1574 TGCCACGCGGTTCCGCTGACCGCTGAAAGCAGCTTGGCCCGACCGGCTTGGCTGCTGCTGCTGCT 1633
1644 CCGCAGCTGCTTTCAGAGACATCTTCTGCGAGAGGGGCGCAGAGGGCTTGGCCAGAGCTG 1703
1634 TCCGTGACGAGCAGCTGSCAGTTACTGATACCACTTTCCGCGATGSCACACAGCTTT 1693
1704 TCGGGAATCACAGGGGCTGCTGCTAATGGACAACTTCCGGGATGCCACAGCTCAC 1763
1694 TGCTTGGACCCGAGTCCGCTCATTTGCACTGAAAGCTGCGGCGAGAGCGCTGCGCAAGC 1753
1764 TACTTGGCACTAGAGTGGGCACACAGCATCTCAAAAAGATTGCACCCCTAGCTTGGCCACA 1823
1754 TGACTCTCTGAGCTTTGTCGCTGAGGCTTGGGCGCGCGACCTACGATGTGGCGATGC 1813
1824 ACTTCAACACCTCTTCAGCATAGAACTGGGAGAGCCACATTTGACGTGGCCATGC 1883
1814 GTTTCCTCTTGGAGATCCGTTGGACAGGCTTCGACGAGCTGCGCGAGCGCATGCCGAATG 1873
1884 CTTCTGTATGATGCTCCCTGGCGGCTTCAGAGCTCCGAGAGCTCATCCCCACA 1943
1874 TAAACATTCAGATGCTGCTCGCGCGCGCAACACCGTGGGATACACCCCGTACCCAGCT 1933
1944 TCCCATTCAGATGCTACTGAGGGGGGCAATGCTGTGGGGCTACACCACTACCCCTGACA 2003
1934 CCGTTCGCGCGCTTTGTTAAGGAAGCTGCCAGCTCCGCGGTGGACATCTTCGCACTC 1993
2004 ACCTGGGCTTCAAGTTCCTGAGGTGGCCAAAGAGAATGGCATGAGCTTCCGGATCT 2063
1994 TCGACGCGCTTAAAGCAGCTCTCCAGATGCTCCAGCAATCGACGAGCTCCTGGAGACCA 2053
2064 TTGACTCCTTAACTACCTGCAAACTGCTGCTGGGCGATGGAAGCAG---CTGGCAGTG 2120
2054 ACACCGCGTACGAGGTGGCTATGCTTATCTGCTGATCTCTGATCCAAATGAAA 2113
2121 CTGGGGGTGGTGGAGCTGCCATCTCTACACGGGTGACGTGGCTGACCCAGTCGCA 2180
2114 AGCTCTACACCTCGATTACTACTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTC 2173
2181 CTAAATACTCACTGGAGTACTACATGGGCTTAGCTGAGAACTGCTGCGAGCGGCACTC 2240
2174 ACATTCTGGCCATTAAAGATATGGCTGCTGCTTCCGCCAGCTGCGGTAACCAAGCTGG 2233

2241 ACATCCTCTGCATTAAAGACATGCGAGSCCTGCTGAAGCCTGCAGCATGCACCTGCTGG 2300
2234 TCACCGCACTGG---CGTGAATTCGATCTGCCAGTGCACGTGCACACCCACGACATG 2290
2301 TCAGTCTCCTCCGGACCGGTTCCCGACCTCCCACTGCACATCCATACCATGACACAT 2360
2291 CGGGTGGCAGTGGCTACCTACTTGTCTGCACTCAAGCTGGTGCAGATGCTGTTGAG 2350
2361 CAGGTCAGGTGGCGACCCATGTTGGCTCTGCAACAGCTGGGCTGATGTTGGAGT 2420
2351 GTGCTTCCGCCACCTGTCTGGCACACCTCCAGCCATCCCTGTCTGCCATGTTGCTG 2410
2421 TGGCAGTGCAGTCTATGCTGGGATGACCTCAGCCCGATGGGGCCCTGGTGGCT 2480
2411 CATTCGCGCACACCCGCTCGCATACCGTGTGAGCCTCGAGGCTGTTCTGACCTCGAGC 2470
2481 GTACCAAGGGGACTCCTCTGGACACAGAGGTACCCCTGGAGCGTGTGTTGACTACAGTG 2540
2471 CGTACTGGGAAGCTGTGCGCGGACTGTACTGCCATTTGA-----GTCTGGAACCCGAG 2524
2541 AGTATTGGGAAGGGCTCGGGGCTGTATGCAGCCTTTGATTGCACGCTACCATGAAGT 2600
2525 GCCCAACCGGTCGCTCTACCGCCACGAAATCCAGGCGGACAGTTGTCACACCTGCGTG 2584
2601 CTGGCACTCAGACGTGTATGAGATGAGATCCAGGGGCGCATACCAACCTACACT 2660
2585 CACAGGCCACCGCACTGGGCTTGTCTGATCGTTGAGCTCATCGAAGACACTACGAC 2644
2661 TCCAGGCCACAGCATGGAGTTGGCTCAAGTTCAAGAGGTCAAGAGGCTATGCTG 2720
2645 CCGTTAATGAGATGTGGGACGCCCAACAGAGTCAACCCATCTCCAAAGTCTGAGTGGG 2704
2721 AGGCTAAACAGATGTGGGGGACCTCATCAAGGTGACACCTCTCCAAAGTGTGGGG 2780
2705 ACCTGCACTCCACTGCTGGTGGTGGGTGATGATCCAGCAGACTTTGCTGCAGACCCAC 2764
2781 ATCTGCGCCAGTTTCATGTCGAAACCGGTTGAGCCGGCGAGGCGAAGCTCAGGCG 2840
2765 AAAAGTACGACATCCAGACTCTGTATCGGCTTCTGCGGGCGAGCTTGGTAAACCTC 2824
2841 AAGAGCTGCTCTCCCGCTCTGTGTGGAGTTCTCTGAGGGCTACATTTGGCATTCGCC 2900
2825 CAGGTGGCTGGCCAGAACACCTGCGCACCCGCGCACTGGAAGGCC 2869
2901 ATGGGGGTTTCCCTGAACCCCTTCCGTTCTAAGTGTCTAAAGGACC 2945

RESULT 15
ABN96532
ID ABN96532 standard; DNA; 4017 BP.

XX AC ABN96532;

XX DT 13-AUG-2002 (first entry)

XX DE Gene #3030 used to diagnose liver cancer.

XX KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
metastatic liver tumour; cytostatic; expression profile; disease state;
disease progression; drug toxicity; drug efficacy; drug metabolism.

XX OS Homo sapiens.

XX PN WO200229103-A2.

XX PD 11-APR-2002.

XX PF 02-OCT-2001; 2001WO-US30589.

XX PR 02-OCT-2000; 2000US-237054P.

XX PA (GENE-) GENE LOGIC INC.

XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX WPI; 2002-426119/45.
XX
XX Diagnosing and detecting the progression of liver cancer,
XX PT hepatocellular carcinoma or metastatic liver tumor in a patient,
XX PT involves detecting the level of expression of two or more genes in a
XX liver tissue sample
XX
XX Claim 1; SEQ ID NO 3030; 298bp; English.
XX
XX The invention relates to a novel method for diagnosing and detecting the
XX progression of liver cancer, hepatocellular carcinoma or metastatic liver
XX tumour in a patient, and differentiating metastatic liver cancer from
XX hepatocellular carcinoma in a patient, involving detecting the level of
XX expression of two or more genes represented in ABN93503-ABN97455 in a
XX tissue sample. The method of the invention has hepatotropic, and
XX cytostatic activity. The method is useful for diagnosing and detecting
XX the progression of liver cancer, hepatocellular carcinoma and metastatic
XX liver carcinoma in a patient. The method is useful for identifying
XX expression profiles which serve as useful diagnostic markers as well as
XX markers that can be used to monitor disease states, disease progression,
XX drug toxicity, drug efficacy and drug metabolism.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 4017 BP; 825 A; 1238 C; 1235 G; 719 T; 0 other;
Query Match 19.4%; Score 674.2; DB 24; Length 4017;
Best Local Similarity 54.3%; Pred. No. 1.2e-173;
Matches 1523; Conservative 0; Mismatches 1258; Indels 24; Gaps 7;
QY 86 TCAAAAGATCTGTAGCAACCGCGGAAATCGCGTCCGCTTTCGTCGACGAC 145
DB 192 TCAGAAAGTCATGTGGCCAAAGAGTGAGATTGCCATCCGTGTTCGCGGCTGCA 251
QY 146 TCGAAACCGGTGCGACCGTAGCTATTTACCCCGGTGAAGATCGGGGATCATTTCCACC 205
DB 252 CGGAGCTGGGCATCCGACCGTAGCCATCTACTCTGACGAGACACGGCCAGATGCACC 311
QY 206 GCTCTTTTGTCTTGAAGCTCTCGCATTTGTGATCTGAAGGCTCACAGTCAAGGGGTACC 265
DB 312 GGCAGAAAGCAGATGAAGCTTATCTATCGCGCCGCGCTGGCCCGCTGTCAGGCTACC 371
QY 266 TGGACATCGATGAATTTATCGTCAGCTAATAAGTTAAGACAGATCTATTACCCGG 325
DB 372 TGCATATCCAGACATCATCAAGGTGGCCAAAGGAGAACACGTAGATGCAATGCACCCCTG 431
QY 326 GATATGGCTTCTGTCTGAAAATGCCAGCTTGCCCGGAGTGGCGGAAACCGCATTA 385
DB 432 GCTACGGGTCTCTCTGAGCGAGGGAGCTTCGCCAGGCTGCGAGATCGAGGGGTCC 491
QY 386 CTTTATTTGCGCCCAACCCAGAGTTCTTGATCTACCGGTGATAAGTCTCGTGGGTAA 445
DB 492 GGTATTATTTGTTCCAAAGCCCAAGTGTGTCGCAAGATGGGAGACAAGTGGAGGCCCGGG 551
QY 446 CCGCGCGAAGAGGCTGTCTGCAGTTTGGCGG---AATCCACCCGAGCAAAACA 502
DB 552 CCATCGCCATTTGTCGGGTGTTCGGGTGTCCTGGCAGATGCCCCCATCAGTCCC 611
QY 503 TCGATGACATCGTTAAAGCGCTGAAGCCAGACTTACCCCATCTTTGTAAAGGCGATTG 562
DB 612 TGCATGAGCCCAACGAGTTCTTCAACACTACGGCTTCCCATCATCTTCAAGGGGCT 671
QY 563 CCGGTGTGCGGACGCGGTATGCGCTTTGTCTTCACTGATGAGCTCCGCAAAATGG 622
DB 672 ATGGGGTGGGGGGCGTGGCATGAGGGTGTGTCACAGCTACGAGGAGCTGGAGGAATT 731
QY 623 CAACAGAAAGCATCTCGTGAAGCTGACGGGCAATTCGGCGACGGTTCGGTATATGTCGAC 682
DB 732 ACACCGGGGCTTACTCAGAGGCTGGCGGCTTTGGGAATGGGGCGCTGTTGTGGAGA 791

QY 683 GTGCTGTGATTAAACCCCGACGACATTTGAAGTGCAGATCTCTTGGCGATCGCACTCGAGAG 742
DB 792 AGTTTCATCGAGAAAGCCACCGGCACATCGAGGTGCAGATCTTTGGGGGACAGTATGGGAACA 851
QY 743 TTGTACACCTTTATGACGTGACTGCTCACTGCGAGCGTCTGTCACCAAAAAGTTCTCGAAA 802
DB 852 TCCTGCACTTGTACGAGAGACTGCTCCATCCAGCGCGGCGCACAGAGGTGTGTCAGA 911
QY 803 TTGCGCCAGCACAGCATTTTGGATCCAGAACTGCGTGTATCGCATTTGTGCGGATCGAGTAA 862
DB 912 TTGCCCCCGCGGCCACCTTGACCCGCGAGCTTCGAGCTCGCTCACCGAGACTCTGTGA 971
QY 863 AGTTCTCCGCTCAITTTGTTTACAGGGCGGGGAAACCGTGAATTTCTTGTGTCGATGAAA 922
DB 972 AACTCGCTAAACAGGTGGGCTACGAGAACCGAGCACCGTGGAGTTCTCTGTGTGACAGGC 1031
QY 923 AGGCAACACAGTTTTCATCGAATTAACACAGTATCCAGTTTGAGCACACCGTGTACTG 982
DB 1032 ACGCAAGCACTACTTTCATCGAGTCACTCCGCCCTGCGAGTGGAGCACACGGTTCACAG 1091
QY 983 AAGAAGTCAACGAGGTGGACCTTGTGAAGCGCGAGATGCGCTTGGCTGCTGTGCAACCT 1042
DB 1092 AGGAGATCACCGAGTAGACTTGTTCATGCTCAGATCCAGTGTGCTGAAGCGAGGACC 1151
QY 1043 TGAAGGAATTTGGTCTGACCCAAAGATTAAGATCAAGACCCAGGTGCGAGCATGTCAGTGC 1102
DB 1152 TACCCGACCTGGGGCTCGCGCAGGAGAACATCCGCATCAACCGGTGTGCCATCCAGTGCC 1211
QY 1103 GCATCACACGAGATCCAAACAAACCGCTTCCGCCAGATACCGGAACATATCACCGGT 1162
DB 1212 GGGTCACACGAGGACCCCGGCCACAGTTTCAGCGGACACCGCGCGCATTTAGGTGT 1271
QY 1163 ACCGCTCACAGGCGGAGCTGGCGTTCGTCT---TGACGGTGCAGCTCAGCTCGGTGGCG 1219
DB 1272 TCCGAGCGGAGGGGATGGGCATCCGCTGGATATGCTTCGCGCTTCCAAGAGGCGG 1331
QY 1220 AAATCACCGCACACTTTGACTTCCATGCTGTGTAATAATGACCTGCGGTGTTCGACTTTG 1279
DB 1332 TCATCTCGCCCACTACGACTCCCTGTGTTCAAGTCAATTCGCCACGCGCAAGAGACCACC 1391
QY 1280 AAATGCTGTGTCGTGCACAGCGCGGTGGCTGAGTTTCAACCGTGTCTGTGTTGCA 1339
DB 1392 CCAGCGCGCCACCAAGATGAGCAGGGCCCTTGCAGGTTCCGTTGTCGAGGTGTGAAGA 1451
QY 1340 CCAACATTTGTTCTTTCGCTGCGTTCGCGGGAAGAGGACTTCACTTTCCAAGCGCATCG 1399
DB 1452 CCAACATCGCTTCTGCGAATGTGTCAACACAGCAGTTCCTTGGCAGGCACTGTGG 1511
QY 1400 CCAACGATTTATCGGCGATCAACCCACCTCTTCAAGGTTCCACCTGCGGATGATGAGC 1459
DB 1512 ACACCGAGTTTCATCGACGAGAACCCAGACGTGTTCCAGCTGCGGCTGACAGAACCGGG 1571
QY 1460 AGGACGCGATCTCGGATTTACTTGGCAGATGTCACCGTGAA-----CAAGCTCATGTTGTG 1515
DB 1572 CCAAAAAGCTGTTGCATCTCGGCCATGTCATGTTAAACGGTCCAAACACCCCGATTC 1631
QY 1516 CGTCCAAAGGATGTTGACGACCAATCGATAAGCTGCCAACATCAAGGATCTGCACTG 1575
DB 1632 CCGTCAAGGCCAGCCCCAGCCCCAGGACCCCGTTGTCCTGCGAGTGCCTATAGGCCGC 1691
QY 1576 CCAGC--CGGTTCCCGTGACCGCTGAAGCAGCTTGGCCCCAGCGCGTTTGTCTGATGATC 1633
DB 1692 CCGCGCTGTTTTCAGAGACATCTCTGTCGAGAGGGGCTGAGGGCTTTTCTCGAGCTG 1751
QY 1634 TCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACTTCCGCGATGTCACACAGTCTT 1693
DB 1752 TGCAGAACCAACCCCGGGCTGCTGTGATGGACAGACCTTCAGGGAGCGCCACAGTCA 1811
QY 1694 TGCTTGCACCCGAGTCCGCTCATTTCCAGCTGAAGCTGGGCGAGAGCCCTTCCCAAGC 1753
DB 1812 TGCTGGCGACTCGTGTGGCGACCCACCATCATCTCAAAAAGATCGCCCCCTATGTTGCCACA 1871

QY 1754 TGA CTCTGAGCTTTTGTCCGTGGAGCCCTGGGCGCGGACCTAGCATGTGGCATGC 1813
Db 1872 ACTTCAGCAAGCTCTTCAGCATGAGAACTGGGAGGAGCCAGTTTGAGTGCCTCATGC 1931
QY 1814 GTTTCCTCTTTGAGGATCCGTGGGACACAGCTCGACGAGCTGGCGAGGCGATCCCAATG 1873
Db 1932 GCTTCCTGTATGATGCGCCCTGGGCGGCGTGCAGAGCTCCGAGAGCTCATCCCCAACA 1991
QY 1874 TAAACATTCAGATCTGCTTCGGGCGCGCAACACCGTGGGATACACCCCGTACCCAGACT 1933
Db 1992 TCCGTTTCCAGATCTGCTGGGGGGGCCAAATGCTGFGGGCTACACCACTACCCAGACA 2051
QY 1934 CCGTCTGCGCGCGTTTGTAAAGGAAGCTGCCAGCTCCGCGGTGGACATCTCCGCACTC 1993
Db 2052 ACGTGGTCTTCAAGTTCGTGAAGTGGCCAAAGAGAAATGGCATGGATGCTTCCGTGTGT 2111
QY 1994 TCGACGGGCTTAAAGAGCTCTCCAGATGGCTCCAGCAATCGACGAGCTCTGGAGACCA 2053
Db 2112 TTGACTCCCTCAACTACTTGGCCCAACATGCTGTGGGCATGGAGGCGG---CAGGAAGTG 2168
QY 2054 ACACGGGCTAGCGAGTGGCTATGGCTTATCTGGTGATCTCTGTATCCAAATGAAA 2113
Db 2169 CCGGAGGCGTGGTGGAGGCTGCCATCTCATACAGGGCGAGCTGGCGGACCCCGCGCA 2228
QY 2114 AGCTCTACCCCTGATTAACCTAAAGATGGCAGAGGAGATCGTCAAGCTCTGGCGCTC 2173
Db 2229 CCAAGTACTCACTGCACTACTACATGGCTTGGCTGAAGCGCTGGTCCGAGCTGGCACCC 2288
QY 2174 ACATTCCTGGGCAATTAAGGATATGGCTGTGCTTCCGCCAGCTGGGTAAACCAAGCTGG 2233
Db 2289 ACATCTGTGATCAGGACATGGCCGGGCTGCTGAAGCCACAGCTGCACCCACGACCGCT 2348
QY 2234 TCACCGCACTCGG---CGGTGAATTCGATCTGCAGTGCAGCTGCACACCCACGACACTG 2290
Db 2349 TCAGCTCCCTCCGGGACCGGCTTCCCGACCTTCCACCTGCACATCCACACCCACGACCGCT 2408
QY 2291 CGGTGGCCAGTTCGGCTACCTACTTCTGCTGAGCTCAAGCTGGTGCAGATGCTGTTGACG 2350
Db 2409 CAGGGGAGGCGTGGGAGCCATGCTGCGCTGTGCCAGGCTGGAGTGATGTTGGGATG 2468
QY 2351 GTGCTTCGCAACCACTGTCTGGCACCACTCCAGCCCATCCCTGTCTGCCATGTTGCTG 2410
Db 2469 TGGCAGCTGATTCATGCTGGGATGACTTCACAGCCAGCATGGGGGCGCTGGTGGCCT 2528
QY 2411 CATTCGGGCAACCCGCTGGGATACCGGTTTGAAGCTCGAGGCTGTTCTGAGCTCGAGC 2470
Db 2529 GTACCAGAGGACTCCCTGGACACAGAGTGGCCATGGAGCGCTGTTTGACTACAGTG 2588
QY 2471 CGTACTGGGAGCTGTGGGAGCTGTACTGCGCATTTGAGTCTGGAAACCCAGGCCCAA 2530
Db 2589 AGTACTGGGAGGGCTCGGGACTGTACCGGCTTTCGACTGCACGGCCACCATGAGT 2648
QY 2531 CCGG-----TCGGCTTACCGCCAGAAATCCGAGGAGAGTTCACACCTGGGTG 2584
Db 2649 CTGGCACTCGGACGTGTATGAATGAGATCCAGGGGGCCAGTACACCACTGCACT 2708
QY 2585 CACAGGCCACCGCACTGGGCTTCTGATCGTTCGAGCTCATCGAAGACAACTACGCGAG 2644
Db 2709 TCCAGGGCCACAGCATGGGGCTTGGCTCCAAAGTTCAAGAGGCTCAAGAGGCTATGTGG 2768
QY 2645 CCGTTAATGAGATCTGGGAGCGCCAAACCAAGGTACCCCATCTCCAGGTTGTTGGCG 2704
Db 2769 AGGCCAACAGATCTGGGCGATCTCATCAAGGTGAGCGCCCTCTCTCAAGATCTGGGGG 2828
QY 2705 ACCTCGCACTCCACTGGTGTGGGTGTAGATCCAGCAGACTTTGTGTGACACCCAC 2764
Db 2829 ACCTGGGCCAGTTTATGTGCAGAAATGGAATGAGCCGGGAGAGGCCGGAAGCTCAGGCGG 2888
QY 2765 AAAAGTACGATCCAGACTCTGTATCGGTTCTCTGCGCGGAGCTTGGTAAACCTC 2824
Db 2889 AAGAGTGTCTTTCCCGCTCCGTTGGTGGAGTTCTCTGACAGGCTACATCGGTGTCCCCC 2948
QY 2825 CAGGTGGCTGGCCAGAACCACTGCGGCACCCCGCACTGGGAAGGCC 2869

Db 2949 ATGGGGTTTCCCGAACCCCTTTCGCTCTAAGGTACTGAAGGACC 2993
Search completed: April 5, 2003, 05:53:14
Job time : 1143 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2003, 07:14:43 ; Search time 297 Seconds
(without alignments)
10260.205 Million cell updates/sec

Title: US-09-974-973-1
Perfect score: 3474
Sequence: 1 95gactgctacccttgg.....tgatcgctgcttctctaa 3474

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 593429 seqs, 438593890 residues

Total number of hits satisfying chosen parameters: 1186958

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3474	100.0	3474	9	US-09-974-973-1
2	3474	100.0	3474	9	US-09-974-973-3
3	3398.8	97.8	3621	9	US-10-045-072-1
4	3344.8	96.3	3420	9	US-09-738-626-765
5	681	19.6	3945	10	US-09-917-800A-1566
6	674.2	19.4	4017	10	US-09-880-107-3029
7	502.8	14.5	3429	10	US-09-815-242-6709
c	488.4	14.1	5030	7	US-08-781-986A-324
9	488.2	14.1	20072	10	US-09-070-927A-89
10	486.8	14.0	3441	10	US-09-815-242-4413
11	428.8	12.3	3222	10	US-09-815-242-8263
12	305	8.8	1362	10	US-09-767-479-7
13	302.2	8.7	1350	10	US-09-815-242-6233
14	297	8.5	3065	10	US-09-767-479-5
15	286	8.2	2238	10	US-09-974-300-2
16	283.8	8.2	1350	10	US-09-815-242-9787
17	259.6	7.5	1347	10	US-09-815-242-7063
18	258.6	7.4	1416	10	US-09-815-242-4160
19	255	7.3	1350	10	US-09-815-242-7965

20	249	7.2	1368	10	US-09-815-242-9519	Sequence 9519, Ap
21	245.8	7.1	1368	10	US-09-815-242-9266	Sequence 9266, Ap
22	243.4	7.0	1773	9	US-09-738-626-3440	Sequence 3440, Ap
23	234.2	6.7	1377	10	US-09-815-242-7224	Sequence 7224, Ap
24	231	6.6	1368	10	US-09-815-242-7461	Sequence 7461, Ap
25	229	6.6	6021	10	US-09-070-927A-458	Sequence 458, App
26	225.8	6.5	1329	10	US-09-815-242-3908	Sequence 3908, Ap
27	225.8	6.5	1371	10	US-09-815-242-6827	Sequence 6827, Ap
28	203.4	5.9	1335	9	US-09-891-641-59	Sequence 59, Appl
29	191.4	5.5	2175	9	US-10-160-501-18	Sequence 18, Appl
30	191.4	5.5	2577	9	US-10-160-501-16	Sequence 16, Appl
c	184	5.3	1359	9	US-09-738-626-777	Sequence 777, App
32	180.8	5.2	989	10	US-09-974-300-5516	Sequence 5516, Ap
33	179.2	5.2	3710	7	US-08-781-986A-405	Sequence 405, App
34	178.8	5.1	1356	10	US-09-815-242-9029	Sequence 9029, Ap
35	167.2	4.8	1346	10	US-09-815-242-4751	Sequence 4751, Ap
c	164.4	4.7	2719	12	US-10-007-693-136	Sequence 136, App
37	163.2	4.7	1338	10	US-09-815-242-4363	Sequence 4363, Ap
38	163.2	4.7	1362	10	US-09-815-242-8464	Sequence 8464, Ap
c	162.8	4.7	2785	7	US-08-781-986A-133	Sequence 133, App
40	162	4.7	2205	9	US-09-938-842A-2078	Sequence 2078, Ap
41	158.4	4.6	1330	10	US-09-974-300-1081	Sequence 1081, Ap
42	143.8	4.1	555	10	US-09-974-300-1108	Sequence 1108, Ap
43	140.2	4.0	1077	10	US-09-815-242-8841	Sequence 8841, Ap
44	119	3.4	546	10	US-09-974-300-1106	Sequence 1106, Ap
45	106.2	3.1	474	10	US-09-815-242-3976	Sequence 3976, Ap

ALIGNMENTS

RESULT 1

US-09-974-973-1
; Sequence 1, Application US/09974973
; Patent No. US20020177202A1
; GENERAL INFORMATION:
; APPLICANT: Hanke, Paul D.
; TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacteri
; FILE REFERENCE: 1533.1230001/MAC/RGM
; CURRENT APPLICATION NUMBER: US/09/974,973
; CURRENT FILING DATE: 2001-10-21
; PRIOR APPLICATION NUMBER: US 60/239,913
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
; LENGTH: 3474
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3474)
US-09-974-973-1

Query Match	100.0%;	Score 3474;	DB 9;	Length 3474;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3474;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	GTGACTGCTATCACCCCTTGGCGGTCTCTCTTGTGAAGGAATAATTACTTGTAGTCTCGACT	60	
Db	1	GTGACTGCTATCACCCCTTGGCGGTCTCTCTTGTGAAGGAATAATTACTTGTAGTCTCGACT	60	
Qy	61	CACACATCTTCAACGCCTTCAGAGATTCAAAAAGATCTTGGTAGCAAAACCGCGGGAATC	120	
Db	61	CACACATCTTCAACGCCTTCAGAGATTCAAAAAGATCTTGGTAGCAAAACCGCGGGAATC	120	
Qy	121	GCGGTCCGTGCTTTCGTCGAGCACTCGAAACCGGTCAGCCAGCTAGTCTATTATACCCC	180	
Db	121	GCGGTCCGTGCTTTCGTCGAGCACTCGAAACCGGTCAGCCAGCTAGTCTATTATACCCC	180	
Qy	181	CGTGAAGATCGGGATCATTTCCACCGCTCTTTGCTTCTGAAGCTGTCCGATGGTACT	240	
Db	181	CGTGAAGATCGGGATCATTTCCACCGCTCTTTGCTTCTGAAGCTGTCCGATGGTACT	240	

Qy	241	GAAGGCTCA	CAAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGACGCTAAAAA	300
Db	241	GAAGGCTCA	CAAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGACGCTAAAAA	300
Qy	301	GTTAAAGCAGATGCTATTATACCGGATATGCTTCTGTCTGAAAAATGCCAGCTTGCC	360	
Db	301	GTTAAAGCAGATGCTATTATACCGGATATGCTTCTGTCTGAAAAATGCCAGCTTGCC	360	
Qy	361	CGCGAGTGC	CGAAAAACGCGATTACTTTATTGCGCCAAACCCAGAGGTTCTTGATCTC	420
Db	361	CGCGAGTGC	CGAAAAACGCGATTACTTTATTGCGCCAAACCCAGAGGTTCTTGATCTC	420
Qy	421	ACCGGTGATAAGTCTCGTGGGTAAACCGCGCGAAGAGGCTGTCTGCAGTTTGGCG	480	
Db	421	ACCGGTGATAAGTCTCGTGGGTAAACCGCGCGAAGAGGCTGTCTGCAGTTTGGCG	480	
Qy	481	GAATCCACCCGAGCAAAACATCGATGACCTTTAAAGCGCTGAAGCGCAGACTTAC	540	
Db	481	GAATCCACCCGAGCAAAACATCGATGACCTTTAAAGCGCTGAAGCGCAGACTTAC	540	
Qy	541	CCCATTTTGTAAAGGAGTTGCGCGTGTGCGGACCGGGTATGCGTTTGTCTTCA	600	
Db	541	CCCATTTTGTAAAGGAGTTGCGCGTGTGCGGACCGGGTATGCGTTTGTCTTCA	600	
Qy	601	CTGATGAGTCCGCAAAATGGCAACAGAAAGCATCTGCGAAGCTGAAGCGGCAATCGCG	660	
Db	601	CTGATGAGTCCGCAAAATGGCAACAGAAAGCATCTGCGAAGCTGAAGCGGCAATCGCG	660	
Qy	661	GACGGTTTCGGTATATGTCGAAACGTGCTGTGATTAAACCCACAGCACATTTGAAGTCAGATC	720	
Db	661	GACGGTTTCGGTATATGTCGAAACGTGCTGTGATTAAACCCACAGCACATTTGAAGTCAGATC	720	
Qy	721	CTTGGCGCATCGCATCGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCACGCT	780	
Db	721	CTTGGCGCATCGCATCGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCACGCT	780	
Qy	781	CGTCAACCAAAAGTTGTCGAAATGCGCCAGCACAGCATTTGGATCCAGACTGCGTGAT	840	
Db	781	CGTCAACCAAAAGTTGTCGAAATGCGCCAGCACAGCATTTGGATCCAGACTGCGTGAT	840	
Qy	841	CGCATTTGTGCGGATGAGTAAAGTTCTGCGCTCGATTGGTTACAGGGCGGGGAACC	900	
Db	841	CGCATTTGTGCGGATGAGTAAAGTTCTGCGCTCGATTGGTTACAGGGCGGGGAACC	900	
Qy	901	GTGGAATCTTGTGCGATGAAAGGCAACACAGTTTTCATCGAAATGAACCCACGTAATC	960	
Db	901	GTGGAATCTTGTGCGATGAAAGGCAACACAGTTTTCATCGAAATGAACCCACGTAATC	960	
Qy	961	CAGGTTGAGCACACCGGTGA	CTGAAGAAGTCAACAGAGTGGACCTGGTGAAGCGCGCAGATG	1020
Db	961	CAGGTTGAGCACACCGGTGA	CTGAAGAAGTCAACAGAGTGGACCTGGTGAAGCGCGCAGATG	1020
Qy	1021	CGCTTTGGCTGCTGGTGC	CAACCTTTGAAGGAATTTGGGTCTGACCCAAAGATAAAGATCAAGAAC	1080
Db	1021	CGCTTTGGCTGCTGGTGC	CAACCTTTGAAGGAATTTGGGTCTGACCCAAAGATAAAGATCAAGAAC	1080
Qy	1081	CACGGTCGACACTGACGTGCGGCATCAACAGGAGAGATCCAAACACCGGCTTCCGCCCA	1140	
Db	1081	CACGGTCGACACTGACGTGCGGCATCAACAGGAGAGATCCAAACACCGGCTTCCGCCCA	1140	
Qy	1141	GATACCGGAATATACCCGGTACCGGTACCGTCAACAGCGGAGCTGGGCTGCTGACGCT	1200	
Db	1141	GATACCGGAATATACCCGGTACCGGTACCGGTCAACAGCGGAGCTGGGCTGCTGACGCT	1200	
Qy	1201	GCAGCTCAGCTCGGTGCGGAATATCAACGACACTTTGACTCATGCTGGTGAATAATGACC	1260	
Db	1201	GCAGCTCAGCTCGGTGCGGAATATCAACGACACTTTGACTCATGCTGGTGAATAATGACC	1260	
Qy	1261	TGCGGTGGTTCCGACTTTGAAACTGCTGTTGCTGCTGCAACAGCGCGCTGCGCTGAGTTC	1320	
Db	1261	TGCGGTGGTTCCGACTTTGAAACTGCTGTTGCTGCTGCAACAGCGCGCTGCGCTGAGTTC	1320	

Qy	1321	ACCGTCTCGGTCTCCACCAACAACTGGTTCTCTGCGTGGTGTCTGCGGAGAGGAC	1338
Db	1321	ACCGTGTCTGGTGTGCAACAACTTGGTTCTTCTGCGTGGTGTCTGCGGAAAGAGCA	1338
Qy	1381	TTCACTTCCAAAGCGCATCGCCACCGGATTTATTCGGCGATACACCAACACTCTCTTCAGGCT	1440
Db	1381	TTCACTTCCAAAGCGCATCGCCACCGGATTTATTCGGCGATACCCACACCTCTCTTCAGGCT	1440
Qy	1441	CCACCTCGCGATGATGACGAGGACGCACTCTCGGATTAATTGGCAGATGTCAACCGTGAAC	1500
Db	1441	CCACCTCGCGATGATGACGAGGACGCACTCTCGGATTAATTGGCAGATGTCAACCGTGAAC	1500
Qy	1501	AAGCCTCATGGTGTGGTCTCAAAGGATGTGCAGACACAAATCGATAAGCTGCCCAACATC	1560
Db	1501	AAGCCTCATGGTGTGGTCTCAAAGGATGTGCAGACACAAATCGATAAGCTGCCCAACATC	1560
Qy	1561	AAGGATCTGCCACTGCGACCGCGTTCCCGTGTACCGCTGAAGACAGCTTGGGCCACCGCGG	1620
Db	1561	AAGGATCTGCCACTGCGACCGCGTTCCCGTGTACCGCTGAAGACAGCTTGGGCCACCGCGG	1620
Qy	1621	TTTGTCTGTGATCTCCGTTCAGCAGACGACACTGGCAGTTACTGATACCACTCTCCGCGAT	1680
Db	1621	TTTGTCTGTGATCTCCGTTCAGCAGACGACACTGGCAGTTACTGATACCACTCTCCGCGAT	1680
Qy	1681	GCACACAGTCTTGTCTTCGACCCGAGTCCGCTCATTTGCGACTGMAAGCTGGCGGCGACCTAC	1740
Db	1681	GCACACAGTCTTGTCTTCGACCCGAGTCCGCTCATTTGCGACTGMAAGCTTCGGCGAG	1740
Qy	1741	GCCTGCGCAAGCTGACTCTTGAGCTTTTTCCTGTTGAGAGCTTGGGCGGCGGACCTAC	1800
Db	1741	GCCTGCGCAAGCTGACTCTTGAGCTTTTTCCTGTTGAGAGCTTGGGCGGCGGACCTAC	1800
Qy	1801	GATGTGGCGATGCGTTTCTCTTTGAGATTCGTTGGGACAGCTCTGCAGAGCTCGGCGAG	1860
Db	1801	GATGTGGCGATGCGTTTCTCTTTGAGGATCCGTTGGGACAGCTCTGCAGAGCTCGGCGAG	1860
Qy	1861	GCATGCGGAATGTAAACATTAGATGCTGCTTTCGCGCCGCAACAACGTTGGGATACAC	1920
Db	1861	GCATGCGGAATGTAAACATTAGATGCTGCTTTCGCGCCGCAACAACGTTGGGATACAC	1920
Qy	1921	CCGTACCCAGACTCCGCTCTGCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC	1980
Db	1921	CCGTACCCAGACTCCGCTCTGCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC	1980
Qy	1981	ATCTTCGCGCATTTTCGACCGCGCTTAAACGACGCTCTCCAGAGTGCCTCAGCAATCGAGCA	2040
Db	1981	ATCTTCGCGCATTTTCGACCGCGCTTAAACGACGCTCTCCAGAGTGCCTCAGCAATCGAGCA	2040
Qy	2041	GTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTATTCTGTTGATCTCTCT	2100
Db	2041	GTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTATTCTGTTGATCTCTCT	2100
Qy	2101	GATCCCAATGAAAGCTCTACACCTCGGATTAACCTTAAAGATGGCAGAGGATCGTC	2160
Db	2101	GATCCCAATGAAAGCTCTACACCTCGGATTAACCTTAAAGATGGCAGAGGATCGTC	2160
Qy	2161	AAGTCTGGCTCACAATCTGGCCATTAAAGATATGGCTGTGTCTGCTTCGCCAGCTGGG	2220
Db	2161	AAGTCTGGCTCACAATCTGGCCATTAAAGATATGGCTGTGTCTGCTTCGCCAGCTGGG	2220
Qy	2221	GTAACCAAGCTGTGTACCGGACGTGGCGGTGAATTCGATCTGCCAGTGCACTGCACACC	2280
Db	2221	GTAACCAAGCTGTGTACCGGACGTGGCGGTGAATTCGATCTGCCAGTGCACTGCACACC	2280
Qy	2281	CAGACACTCGGGTGGCCAGTTGGCTACCTACTTTGCTGCAAGTCTAAAGCTGGTGCAAT	2340
Db	2281	CAGACACTCGGGTGGCCAGTTGGCTACCTACTTTGCTGCAAGTCTAAAGCTGGTGCAAT	2340
Qy	2341	GCTGTGACGGTCTTCCGACCAACCTGTCTGCACCACTCCAGGCCATCCCTGCTGTGCC	2400
Db	2341	GCTGTGACGGTCTTCCGACCAACCTGTCTGCACCACTCCAGGCCATCCCTGCTGTGCC	2400
Qy	2401	ATTGTTGCTGCAATTCGCGGACACCCCGTCCGATACCGGTTTGTAGCCTTCGAGGCTGTTTCT	2460

```
Db 2401 ATTGTTGCTGATTCGGCGCACACCGTCGCGATACCGGTTTGAAGCCTCGAGGCTGTTTCT 2460
Qy 2461 GACCTCGAGCGGTACTGGGAAGCTGTGCGGGAAGTGTACCTGCCATTTGAGTCTGGAAAC 2520
Db 2461 GACCTCGAGCGGTACTGGGAAGCTGTGCGGGAAGTGTACCTGCCATTTGAGTCTGGAAAC 2520
Qy 2521 CCAGGCCCAACCGGTGCGGTCTACCGGCAAGAAATCCAGGCGGACAGTGTGTCACACCTG 2580
Db 2521 CCAGGCCCAACCGGTGCGGTCTACCGGCAAGAAATCCAGGCGGACAGTGTGTCACACCTG 2580
Qy 2581 CGTCACAGGCCACCGGCTGCGGTCTGATCGCTTCCAGGCTCATCGAAGCAACTAC 2640
Db 2581 CGTCACAGGCCACCGGCTGCGGTCTGATCGCTTCCAGGCTCATCGAAGCAACTAC 2640
Qy 2641 GCAGCCGTTAATGAGATGCTGGGAGCGGCCCAACCAAGGTCACCCCATCTCCCAAGGTTGTT 2700
Db 2641 GCAGCCGTTAATGAGATGCTGGGAGCGGCCCAACCAAGGTCACCCCATCTCCCAAGGTTGTT 2700
Qy 2701 GCGACCTCGACATCCACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2760
Db 2701 GCGACCTCGACATCCACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2760
Qy 2761 CCACAAAGTACGACATCCACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2820
Db 2761 CCACAAAGTACGACATCCACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2820
Qy 2821 CCTCCAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2880
Db 2821 CCTCCAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2880
Qy 2881 AAGGCACTCTGACGGAAGTTCCTGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940
Db 2881 AAGGCACTCTGACGGAAGTTCCTGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940
Qy 2941 AAGGAAGCTGCGACAGCTCAACGGCTGCTGTTCCGGAAGCCAAACGGAAGGTTCTCTC 3000
Db 2941 AAGGAAGCTGCGACAGCTCAACGGCTGCTGTTCCGGAAGCCAAACGGAAGGTTCTCTC 3000
Qy 3001 GAGCACCGTCCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATCTTCTACGGA 3060
Db 3001 GAGCACCGTCCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATCTTCTACGGA 3060
Qy 3061 CTGCTGAGGCGCGGAGACTTTGATCCGCTGCCAGATGCGCACCCCACTGCTTGT 3120
Db 3061 CTGCTGAGGCGCGGAGACTTTGATCCGCTGCCAGATGCGCACCCCACTGCTTGT 3120
Qy 3121 CGCTGGATGCGATCTCTGAGCCAGAGAGATAAGGGTATGCGCAATGTTGTGGCCAAAGTC 3180
Db 3121 CGCTGGATGCGATCTCTGAGCCAGAGAGATAAGGGTATGCGCAATGTTGTGGCCAAAGTC 3180
Qy 3181 AACGGCCAGATCCGCCCAATGCGTGTGCGTACCGCTCCGTTGAGTGTGTCACCGCAACC 3240
Db 3181 AACGGCCAGATCCGCCCAATGCGTGTGCGTACCGCTCCGTTGAGTGTGTCACCGCAACC 3240
Qy 3241 GCAGAAAAGGCGAGATCTCTCAACAAAGGGCCATGTTGCTGACCACTTCGCTGGTGTGTC 3300
Db 3241 GCAGAAAAGGCGAGATCTCTCAACAAAGGGCCATGTTGCTGACCACTTCGCTGGTGTGTC 3300
Qy 3301 ACTGTGATGTTGCTGAAGGTGATGAGGTGAGGTCGAGTGCAGTGCAGTGCAGTGCAGTGCAG 3360
Db 3301 ACTGTGATGTTGCTGAAGGTGATGAGGTGAGGTCGAGTGCAGTGCAGTGCAGTGCAGTGCAG 3360
Qy 3361 GCTATGAAGTGAAGCAACATCAGTCTCTGTCGACGCAAGATGGAACGGTGTG 3420
Db 3361 GCTATGAAGTGAAGCAACATCAGTCTCTGTCGACGCAAGATGGAACGGTGTG 3420
Qy 3421 GTTCTGCTCAACGAAGGTGGAAGGTGGAAGGTGGAAGGTGGAAGGTGGAAGGTGGAAGGTGGA 3474
Db 3421 GTTCTGCTCAACGAAGGTGGAAGGTGGAAGGTGGAAGGTGGAAGGTGGAAGGTGGAAGGTGGA 3474
```

RESULT 2

```
US-09-974-973-3
; Sequence 3, Application US/09974973
; Patent No. US20020177202A1
; GENERAL INFORMATION:
; APPLICANT: Hanke, Paul D.
; TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacteri
; FILE REFERENCE: 1533.1230001/MAC/RGM
; CURRENT APPLICATION NUMBER: US/09/974,973
; PRIOR FILING DATE: 2001-10-21
; PRIOR APPLICATION NUMBER: US 60/239,913
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 3474
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-974-973-3
```

```
Query Match 100.0%; Score 3474; DB 9; Length 3474;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 GTGACTGCTATCACCCCTTGGGGGTCTCTTGTGAAAGGAATAATTACTCTAGTGTGCACT 60
Db 1 GTGACTGCTATCACCCCTTGGGGGTCTCTTGTGAAAGGAATAATTACTCTAGTGTGCACT 60
Qy 61 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAAACCGCGGCGAAATC 120
Db 61 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAAACCGCGGCGAAATC 120
Qy 121 GCGGTCCGTCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGCTAGCTATTATACCCC 180
Db 121 GCGGTCCGTCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGCTAGCTATTATACCCC 180
Qy 181 CGTGAAGATCGGGATCATTCACGCTCTTTGCTTCTGAAGCTGTCCGATTTGGTACT 240
Db 181 CGTGAAGATCGGGATCATTCACGCTCTTTGCTTCTGAAGCTGTCCGATTTGGTACT 240
Qy 241 GAAGGCTCACAGTCAAGGGGTACCTGGACATCGATGAAATTTACGGTGCAGCTTAAAAA 300
Db 241 GAAGGCTCACAGTCAAGGGGTACCTGGACATCGATGAAATTTACGGTGCAGCTTAAAAA 300
Qy 301 GTTAAAGCAGATGCTATTTACCCGGGATATGGCTTCTGTGAAAAATGCCAGCTTGCC 360
Db 301 GTTAAAGCAGATGCTATTTACCCGGGATATGGCTTCTGTGAAAAATGCCAGCTTGCC 360
Qy 361 CGCGAGTCCGGAAGCAACCGCATTTATTTATTTGGCCCAACCCAGAGGTTCTTGATCTC 420
Db 361 CGCGAGTCCGGAAGCAACCGCATTTATTTATTTGGCCCAACCCAGAGGTTCTTGATCTC 420
Qy 421 ACCGTTGATGAAGTCTCGTGGGTAAACCGCCGGAAGAGCTGGTCTGCCAGTTTTCGCG 480
Db 421 ACCGTTGATGAAGTCTCGTGGGTAAACCGCCGGAAGAGCTGGTCTGCCAGTTTTCGCG 480
Qy 481 GAATCCACCCCGAGCAAAAACATCGATGACATCGTGTAAAAAGCGCTGAAGCCAGACTTAC 540
Db 481 GAATCCACCCCGAGCAAAAACATCGATGACATCGTGTAAAAAGCGCTGAAGCCAGACTTAC 540
Qy 541 CCATCTTTGTAAGGAGTTGCGGTGGTGGCGGACCGCGGTATGCGCTTTGTTCTTCA 600
Db 541 CCATCTTTGTAAGGAGTTGCGGTGGTGGCGGACCGCGGTATGCGCTTTGTTCTTCA 600
Qy 601 CTTGATGAGTCCGCAAAATGGCAACAGAGCATCTCGTGAAGCTGAAGCGGATTCGGC 660
Db 601 CTTGATGAGTCCGCAAAATGGCAACAGAGCATCTCGTGAAGCTGAAGCGGATTCGGC 660
Qy 661 GACGGTTGCGTATATGTCGAAACGCTGTGATTAACCCCGCAGCACATTTGAAGTCAGATC 720
Db 661 GACGGTTGCGTATATGTCGAAACGCTGTGATTAACCCCGCAGCACATTTGAAGTCAGATC 720
Qy 721 CTTGGGATCGCACTGGAGAGTTGTACACCTTTATGAACGTGACTGCTCACTGACGCT 780
Db 721 CTTGGGATCGCACTGGAGAGTTGTACACCTTTATGAACGTGACTGCTCACTGACGCT 780
```


QY 3361 GCTATGAAGATGGAGCAACAATCACTCTCTCTTGAACGGAGATTGAACGGTTGTG 3420
DB 3508 GCTATGAAGATGGAGCAACAATCACTCTCTCTTGAACGGAGATTGAACGGTTGTG 3567
QY 3421 GTTCTCTGCTCAACGAAGGTGGAGGTGGCGACTTGTATCGTCGTCGTTCTCTAA 3474
DB 3568 GTTCTCTGCTCAACGAAGGTGGAGGTGGCGACTTGTATCGTCGTCGTTCTCTAA 3621

RESULT 4
US-09-738-626-765
; Sequence 765, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, WASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 765
; LENGTH: 3420
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-765

Query Match 96.3%; Score 3344.8; DB 9; Length 3420;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3373; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 52 GTGTCGACTCACACATCTTCAACGCTTCCAGCATTTCAAAAGATCTTGGTAGCAAAACCGC 111
DB 1 GTGTCGACTCACACATCTTCAACGCTTCCAGCATTTCAAAAGATCTTGGTAGCAAAACCGC 60

QY 112 GGGGAAATCGCGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCT 171
DB 61 GGGGAAATCGCGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCT 120

QY 172 ATTATACCCCGTGAAGATCCGGGATCATTCACCGCTCTTTTGTCTTGAAGCTGTCCGC 231
DB 121 ATTATACCCCGTGAAGATCCGGGATCATTCACCGCTCTTTTGTCTTGAAGCTGTCCGC 180

QY 232 ATTGGTACTGAAGCTCACAGTCAAGCGGTACTCGACATCGATGAATATCGGTGCA 291
DB 181 ATTGGTACCAGGCTCACAGTCAAGCGGTACTCGACATCGATGAATATCGGTGCA 240

QY 292 GCTAAAAAGTTAAAGCAGATGCTATTATCCCGGATATGGCTTCCTGTCTGAAATGCC 351
DB 241 GCTAAAAAGTTAAAGCAGATGCCATTATCCCGGATATGGCTTCCTGTCTGAAATGCC 300

QY 352 CAGCTTCCCGAGTCCGCGGAAACCGCATTAATTTTATTTGGCCCAACCCAGAGGTT 411
DB 301 CAGCTTCCCGAGTCCGCGGAAACCGCATTAATTTTATTTGGCCCAACCCAGAGGTT 360

QY 412 CTTGATCTACCCGTTGATAGTCTCGGTGCGGTAAACCGCGGAGAGGCTGTCTGCCA 471
DB 361 CTTGATCTACCCGTTGATAGTCTCGGTGCGGTAAACCGCGGAGAGGCTGTCTGCCA 420

QY 472 GTTTTGGCGGAATCCACCCCGAGCAAAAAACATCGATGACATCGTTTAAAGCGCTGAAGGC 531
DB 421 GTTTTGGCGGAATCCACCCCGAGCAAAAAACATCGATGATCGTTTAAAGCGCTGAAGGC 480

QY 532 CAGACTTACCCCATCTTTGTAAAGGAGTTTCCGCTGGTGGCGGACGCGGTATGCGCTTT 591
DB 481 CAGACTTACCCCATCTTTGTAAAGGAGTTTCCGCTGGTGGCGGACGCGGTATGCGCTTT 540

QY 592 GTTTTCTTACCTGATGAGCTCCGCAAAATTTGGCAACAGAAGCATCTCTGTAAGCTGAAGCG 651
DB 541 GTTTGCTTACCTGATGAGCTTCCGCAAAATTTAGCAACAGAAGCATCTCTGTAAGCTGAAGCG 600

QY 652 GATTCGGCGACGGTTCCGTTATGTCGAACGTCGTGTGATTAACCCCCAGACATTTGAA 711
DB 601 GCTTTCGGCGATGGCGCGGTATATGTCGAACGTCGTGTGATTAACCCCTCAGCATTTGAA 660

QY 712 GTCAGATCCTTTCGCGATCGACCTGGAGAGTTGTACACCTTTATGAACGTGACTGCTCA 771
DB 661 GTCAGATCCTTTCGCGATCACCTGGAGAGTTGTACACCTTTATGAACGTGACTGCTCA 720

QY 772 CTGACGCTGCTCACCAAAAAGTTGTCGAAATTTGCGCCAGCACAGCATTTTGGATCCAGAA 831
DB 721 CTGACGCTGCTCACCAAAAAGTTGTCGAAATTTGCGCCAGCACAGCATTTTGGATCCAGAA 780

QY 832 CTCGCTGATCGCATTTTTCGCGATGAGTAAAGTTCTGCGCTCCATTTGTTTACAGGGC 891
DB 781 CTCGCTGATCGCATTTTTCGCGATGAGTAAAGTTCTGCGCTCCATTTGTTTACAGGGC 840

QY 892 GCGGGAACCGTGGAAATCTTTGGTGGATGAAAGGGCAACACCGTTTTCATCGAAATGAAC 951
DB 841 GCGGGAACCGTGGAAATCTTTGGTGGATGAAAGGGCAACACCGTTTTCATCGAAATGAAC 900

QY 952 CCACGTATCCAGGTGAGCACACCGTGACTGAAGAGTACCAGAGTGACCTGGTGAAG 1011
DB 901 CCAGGTATCCAGGTGAGCACACCGTGACTGAAGAGTACCAGAGTGAGCTGGTGAAG 960

QY 1012 GCGCAGATGCGTTTGGTCTGCTGTCGCAACCTTGAAGGAATTTGGTCTGACCCAGATAAG 1071
DB 961 GCGCAGATGCGTTTGGTCTGCTGTCGCAACCTTGAAGGAATTTGGTCTGACCCAGATAAG 1020

QY 1072 ATCAAGACCCAGGTGCAGCACTGCGATGCGCATCACACGGAGAGATCCAAACAGCGC 1131
DB 1021 ATCAAGACCCAGGTGCAGCACTGCGATGCGCATCACACGGAGAGATCCAAACAGCGC 1080

QY 1132 TTCGCGCCAGATACCGGAATATCACCGGTACCGCTCACAGCGGAGCTGGCGTTCTGT 1191
DB 1081 TTCGCGCCAGATACCGGAATATCACCGGTACCGCTCACAGCGGAGCTGGCGTTCTGT 1140

QY 1192 CTTGACCGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTTGACTCCATGCTGGTG 1251
DB 1141 CTTGACCGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTTGACTCCATGCTGGTG 1200

QY 1252 AAATGACCTGCGTGGTCCGACTTTGAACTGCTGTGCTGTCGACAGCGCGCGTTG 1311
DB 1201 AAATGACCTGCGTGGTCCGACTTTGAACTGCTGTGCTGTCGACAGCGCGCGTTG 1260

QY 1312 GCTGAGTTTCAACCGTCTGCTGTTGCAACCAATTTGGTCTTCTGCTGCTGCTGCGG 1371
DB 1261 GCTGAGTTTCAACCGTCTGCTGTTGCAACCAATTTGGTCTTCTGCTGCTGCTGCGG 1320

QY 1372 GAAGAGGACTTCACTTCCAAAGCGCATCGCCACCGGATTTATCGGCGGATCCACACACTC 1431
DB 1321 GAAGAGGACTTCACTTCCAAAGCGCATCGCCACCGGATTTATCGGCGGATCCACACACTC 1380

QY 1432 CTTGACGCTCCACTCGGATGATGAGAGGAGCGATCCTGGATTTACTTTGGCAGATGTC 1491
DB 1381 CTTGACGCTCCACTCGGATGATGAGAGGAGCGATCCTGGATTTACTTTGGCAGATGTC 1440

QY 1492 ACCGTGAACAAAGCCTCATGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1551
DB 1441 ACCGTGAACAAAGCCTCATGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1500

1552 CCCAACATCAAGGATCTGCCACTGCCACGCGTTCCCGTGACCGCTGAAGCAGCTTGGC 1611
1501 CCTAACATCAAGGATCTGCCACTGCCACGCGTTCCCGTGACCGCTGAAGCAGCTTGGC 1560
1612 CCAGCGCGTTTGGTCTGGATCTCCGTCAGCAGGAGCGACTGGCATTTACTGATACACC 1671
1561 CCAGCGCGTTTGGTCTGGATCTCCGTCAGCAGGAGCGACTGGCATTTACTGATACACC 1620
1672 TTCCGCGATGACACACGATCTTTGCTTGGACCGCGAGTCCGCTCATTTCCGACTGAAGCCT 1731
1621 TTCCGCGATGACACACGATCTTTGCTTGGACCGCGAGTCCGCTCATTTCCGACTGAAGCCT 1680
1732 GCGCAGAGCGCGTCGCAAGAGTGAATCTTCCGAGCTTTTTCGTCGTCGAGGCTCGGCGGC 1791
1681 GCGCAGAGCGCGTCGCAAGAGTGAATCTTCCGAGCTTTTTCGTCGTCGAGGCTCGGCGGC 1740
1792 GCGACCTACGATGTGGCGATGCGTTTCCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAG 1851
1741 GCGACCTACGATGTGGCGATGCGTTTCCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAG 1800
1852 CTGCGCGAGCGGATGCGCAATGTAAACATTCAGATGCTGCTTCGCGCGCGCAACACCGTG 1911
1801 CTGCGCGAGCGGATGCGCAATGTAAACATTCAGATGCTGCTTCGCGCGCGCAACACCGTG 1860
1912 GGATACACCCCGTACCCAGACTCCGTCCTGCGCGCGTTTCTTAAGGAAGCTGCCAGCTCC 1971
1861 GGATACACCCCGTACCCAGACTCCGTCCTGCGCGCGTTTCTTAAGGAAGCTGCCAGCTCC 1920
1972 GCGGTGACATCTTCGCGATCTTCGACGCGCTTAAACGAGCTCTCCAGATGCGTCCAGCA 2031
1921 GCGGTGACATCTTCGCGATCTTCGACGCGCTTAAACGAGCTCTCCAGATGCGTCCAGCA 1980
2032 ATCGACGAGCTCTCGAGACCAACACCGCGGTAGCGAGGTGCTATGGCTTATTCGTGT 2091
1981 ATCGACGAGCTCTCGAGACCAACACCGCGGTAGCGAGGTGCTATGGCTTATTCGTGT 2040
2092 GATCTCTGATCCAAATGAAGCTCTACACCTCGGATTAATACTAAAGATGGCAGAG 2151
2041 GATCTCTGATCCAAATGAAGCTCTACACCTCGGATTAATACTAAAGATGGCAGAG 2100
2152 GAGATCGTCAAGTCTGCGGCTCAGATTTCTGGCCATTAAGGATATGGTGTCTGCTTCGC 2211
2101 GAGATCGTCAAGTCTGCGGCTCAGATTTCTGGCCATTAAGGATATGGTGTCTGCTTCGC 2160
2212 CCAGCTCGGTAAACCAAGCTGGTCAACCGCACTGCGCGCTGAAATTCGATCTGCCAGTGCAC 2271
2161 CCAGCTCGGTAAACCAAGCTGGTCAACCGCACTGCGCGCTGAAATTCGATCTGCCAGTGCAC 2220
2272 GTGCACACCCAGCAGACTGCGGGTGCGCAGTTGCTACTTTGCTGCGAGCTCAAGCT 2331
2221 GTGCACACCCAGCAGACTGCGGGTGCGCAGTTGCTACTTTGCTGCGAGCTCAAGCT 2280
2332 GGTGCAGATGCTGTGAGGTGCTTCGCGACCACTGCTCGGACCACTCCAGCCATCC 2391
2281 GGTGCAGATGCTGTGAGGTGCTTCGCGACCACTGCTCGGACCACTCCAGCCATCC 2340
2392 CTGTCTGCCATTTGCTGCATTCGCGACACCGCGTCCGATACCGGTTTGGCCTCGAG 2451
2341 CTGTCTGCCATTTGCTGCATTCGCGACACCGCGTCCGATACCGGTTTGGCCTCGAG 2400
2452 GCTGTTTTCTGACCTCGAGCGCTACTGCGAAGCTGTGCGCGAGCTGTACTGCCATTGAG 2511
2401 GCTGTTTTCTGACCTCGAGCGCTACTGCGAAGCTGTGCGCGAGCTGTACTGCCATTGAG 2460
2512 TCTGGAACCCCGAGCCCAACCGGTCGCGTCTACCGGCAACGAAATCCAGCGGACAGATTG 2571
2461 TCTGGAACCCCGAGCCCAACCGGTCGCGTCTACCGGCAACGAAATCCAGCGGACAGATTG 2520
2572 TCCAACTCGGTGACAGGCCAGCCAGCTGGGCTTCTGATCGCTTCGAGCTCATCGAA 2631
2521 TCCAACTCGGTGACAGGCCAGCCAGCTGGGCTTCTGATCGCTTCGAGCTCATCGAA 2580
2632 GACAACTACGCGCGGTTAATGAGATCTGGGACGCCCAACCAAGGTCACCCCATCTCC 2691

2581 GACAACTACGCGCGGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCTCC 2640
2692 AAGGTTGTTGGCGACTCGCACTCCACCTGTTGGTGGGGTGTAGATCCAGCAGACTTT 2751
2641 AAGGTTGTTGGCGACTCGCACTCCACCTGTTGGTGGGGTGTAGATCCAGCAGACTTT 2700
2752 GCTGCAGACCCACAAAGATGACATCCAGACTCTGTCTCATCGGTTCTCGGGCGGAG 2811
2701 GCTGCCGATGACAAAGATGACATCCAGACTCTGTCTCATCGGTTCTCGGGCGGAG 2760
2812 CTTGTGTAACCTTCAGGTGGCTGGCCAGAACCACTGCGCCACCCCGGCACTCGAAGCGCGC 2871
2761 CTTGTGTAACCTTCAGGTGGCTGGCCAGAGCACTGCGCAACCCGCGCACTGGAAGCGCGC 2820
2872 TCCGAAGCAAGGCACTCTGACCGAAGTTCTGAGGAAGAGCGGCGCACTCGACGCT 2931
2821 TCCGAAGCAAGGCACTCTGACCGAAGTTCTGAGGAAGAGCGGCGCACTCGACGCT 2880
2932 GATGATTCGAAGGAACGTCGCAACAGCCTCAACCGCTGCTGTTCCCGAAGCCAAACCGAA 2991
2881 GATGATTCGAAGGAACGTCGCAATAGCCTCAACCGCTGCTGTTCCCGAAGCCAAACCGAA 2940
2992 GAGTTCTCTGAGCAGCCGTCGCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTC 3051
2941 GAGTTCTCTGAGCAGCCGTCGCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTC 3000
3052 TTCTACGCACTGCTCGAGGCGCGGAGACTTTGATCCGCTGCCAGATGTGGCACCCCA 3111
3001 TTCTACGCACTGCTCGAGGCGCGGAGACTTTGATCCGCTGCCAGATGTGGCACCCCA 3060
3112 CTGCTGTTGCTGCTGGATGCGATCTCTGAGCAGACGATTAAGGATTCGGAATTTGTG 3171
3061 CTGCTGTTGCTGCTGGATGCGATCTCTGAGCAGACGATTAAGGATTCGGAATTTGTG 3120
3172 GCCAAGCTCAACGCGCAGATCCGCCAATGCGTGTGCGTGAACCGCTCCGTTGAGTCTGTC 3231
3121 GCCAAGCTCAACGCGCAGATCCGCCAATGCGTGTGCGTGAACCGCTCCGTTGAGTCTGTC 3180
3232 ACCGCAACCGCAGAAAGCGCAGATTCTCCCAACAGGCGCATGTTGTCGCCAATTCGCT 3291
3181 ACCGCAACCGCAGAAAGCGCAGATTCTCCCAACAGGCGCATGTTGTCGCCAATTCGCT 3240
3292 GGTGTTGTCACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCGATCGCA 3351
3241 GGTGTTGTCACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCGATCGCA 3300
3352 ATCATCGAGGCTATCAAGATGGAACCAACATCACTGCTTCTGTTGACGCGCAAGATTGAA 3411
3301 ATCATCGAGGCTATCAAGATGGAACCAACATCACTGCTTCTGTTGACGCGCAAAATCGAT 3360
3412 GCGTGTGTTGCTGCTGCTGCAACGAGGTGGAAGTGGGACTTGATCGTGGTTCCTCC 3471
3361 GCGTGTGTTGCTGCTGCTGCAACGAGGTGGAAGTGGGACTTGATCGTGGTTCCTCC 3420

RESULT 5

US-09-917-800A-1566
; Sequence 1566, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31

; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1566
; LENGTH: 3945
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NW_012744
US-09-917-800A-1566

Query Match 19.6%; Score 681; DB 10; Length 3945;
Best Local Similarity 54.2%; Pred. No. 1.3e-200;
Matches 1521; Conservative 0; Mismatches 1260; Indels 24; Gaps 6;

QY 86 TCAAAAGATCTTGGTAGCAACCGCGGCGAATCGCGGTCCGCTTTCGGTGCAGCAC 145
DB 144 TCAAGAAAGTAATGGTGGCGCAACAGAGGTGAGATTGCCATCCGAGTGTTCGTGCTGCA 203

QY 146 TCGAAACCGGTGCAGCAGGTAGCTATTACCCCGTGAAGATCGGGATCATTCACACC 205
DB 204 CAGAGCTGGGTATCCGACAGTGGCTGTCTACTCGAGCAGGACACAGGCCAGATGCACC 263

QY 206 GCTCTTTTGGCTTCTGAAGCTGTCCGCAATTTGGTACTGAAGGCTCACCAAGTCAAGCGGTACC 265
DB 264 GGCAGAAAGCTGATGAAGCTTACCTTATTTGGCCGTGGCTCCCTGTCGAAAGCTTACC 323

QY 266 TGGACATCGATGAATATTCGGTGCAGCTAAAGTAAAGTTAAGCAGATGCTATTATTCACCGG 325
DB 324 TGCAATTTCCAGACATCATTAAGGTGGCCAGGAGAAATGGTGTAGATGCTGTGCACCCCTG 383

QY 326 GATATGCTTCTCTGTAAGATCCCGAGCTTGCCTCGAGTGCAGGAGTGCAGGAAACCGCATTA 385
DB 384 GCTATGGGTTCTCTCAGAGAGCAGACTTTGCCAGGCTGCCAAGATGCTGGAGTCC 443

QY 386 CTTTATTGGCCCAACCCAGAGGTTCTTGATCTTCACCGGTGATAAGTCTCGTCCGTTAA 445
DB 444 GATTTCATTGGTCCAAAGCCAGAGGTGTCCGCAAGATGGGAGACAAAGTGGAAAGCCCGG 503

QY 446 CCGCCGGAAGAGGCTGGTCTGCAGTTTGGCGGAATCCA-----CCCGAGCAAAAACA 502
DB 504 CCATTTGCATTTGCTGCAGGCGTTCAGTGGTCCCTGGCACTAATTTCCCGCCCATCAATTC 563

QY 503 TCGATGACATCGTTAAAGCGCTGAAGCCAGACTTACCCCATCTTTGTAAAGCAGTTG 562
DB 564 TGCATAGGCACACGATTTCTCTAACACCTATGTTTTCCTTATTTATCTTCAAGGCTGCCT 623

QY 563 CCGTGTGTGGCGGACGCGGTATGCGCTTTGTTTCTTCCACCTGATGAGCTCCGCAAAATGG 622
DB 624 ATGGAGTGGGGCGGTGGCATGAGGTGTGTGCATAGCTACGAGGAGCTGGAGAGATTT 683

QY 623 CAACAGAAAGCATCTGTGAAGCTGAAGCGGATTCGCGCAGGTTTGGGTATATGTGGAAC 682
DB 684 ACACCCGGGCTACCTTGAAGGCTTGGCAGCTTTGGGAATGGGGCATGTTTGTGGAGA 743

QY 683 GTGCTGTGATTAAACCCAGCAGATTAAGTGCAGATCTTGGGATCGCACTGGAGAAG 742
DB 744 AATTCAATTGAAGCCAAAGACACATTGAGGTGCAGATCTTAGGGGACCAATATATGGGAACA 803

QY 743 TTGTACACCTTTTATGAACGTGACTGCTCACTGCAGCGTCTGTCACCAAAAAGTTGTGAAA 802
DB 804 TCTTGCACTTGTATAGCGGGACTGCTCCATCCAGCGCGGCACAGAAAGTGTGTAGAGA 863

QY 803 TTGGCCAGCAGCAGCATTGATCCAGAACTGGGTGATCCGATTTGTGCGGATGCAGTAA 862
DB 864 TTGCCCCCTGTACCCACCTGGACCCCACTTCCGCTCACGCTCACCAAGTACTCTGTCA 923

QY 863 AGTTTCTGCCCTCCATTTGGTTTACAGGGCGCGGAAACCGTGGAAATCTTGGTCATGAAA 922
DB 924 AACTTGCACAGCAGGTGGCTATGAGNATCGAGCACTGTGGAGTTCTTGGTGACAAAGC 983

QY 923 AGGCAACACAGTTTTCATCGAAATGAACCCAGTATCCAGGTGAGCACAACCGTACTG 982
DB 984 ATGCAAGCACTACTTTCATCGAGTCAATTTCCGCTGTCAGGTGGAGCACAACCGTCACTG 1043

QY 983 AAGAAGTCCAGGAGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCT 1042
DB 1044 AGGAGATTACAGATGTGGACCTGGTCCATGCTCAGATCCATGTGTCCGAAGCGCGAGCC 1103

QY 1043 TGAAGGAATTTGGTCTGACCCAAAGATAAGATCAAGACCCACGCTGCAGCACTGCAGTGCC 1102
DB 1104 TGCTGACCTAGGCTTGGCGCAGGAAAATCCGATCAATGTTGTGCCATTCAGTGTG 1163

QY 1103 GCATCACCGAAGATCCAAACACCGGTTCCGCCAGATACCGGAACTATCACCCGCT 1162
DB 1164 GGGTCACCACTAGGACCTTGCAGCAGCTTCCAGCCAGACACTGGCCGCAATTCAGGTTT 1223

QY 1163 ACCGTCACAGCGGAGCTGGGCTTGTGTTGACGGT---CAGTCACTCGGTGGCG 1219
DB 1224 TCCGAGTGGTGGGCGATGGGCTCCGCTGGACAAATGCTCAGCATTCAGGAGCTG 1283

QY 1220 AAATCACCGCACACTTTGACTCCATGCTGTGAAATGACCTGCGGTGTTCCGACTTTG 1279
DB 1284 TCATATCCCCHACTATGACTCCTGCTGTCAAAGTCAATGCTCCATGGCAAGACCAAC 1343

QY 1280 AAATGCTGTGTGTCAGCGCGGTTGGCTGAGTTTCACTGCTGTGTTGTTGCTGCTG 1339
DB 1344 CTACAGCTGCCACCAAGATGAGCAGAGCCCTGGCGGAGTTCCGTGTCGAGGTGTAAGA 1403

QY 1340 CCAACATTTGTTTTCGTGCTGCTGCGGGAAGAGGACTTCACTTCAAGGGCATCG 1399
DB 1404 CCAACATCCCTTCTGCAAGATGTGCTCAAACACGAGAGTTCTTAGCGGGCATTTGG 1463

QY 1400 CCACCGATTTATCGCGATCACCACTCTTCAAGGCTCCAGCTGCGGATGATGAGC 1459
DB 1464 ACACCGATTCATCGATGAGAACCCGAGCTGTTCCAGCTGGCGCTGCAAGAACCGGG 1523

QY 1460 AGGACCGCATCTGGATTACTTGGCAGATGTACCGGTGAACAAGCCCTCATGTTGCTG 1519
DB 1524 CCAGAAAGTTGTACATTACTTGGACAGCTCATGTTCAATGGCCCTTACCCTCCAATCC 1583

QY 1520 CAAAGATTTGAGCAGCAACCATCGATAAGC-----TGCCCAACATCAAGGATTCGCCAC 1573
DB 1584 CCGTCAAGGTCAGTCCAGCCCTGTGGACCCCATTTGCTGTGTTGGTCCCATAGGCCAC 1643

QY 1574 TGCCACGCGTTCCCGTGACCGCTGAAGCAGCTTGCSCCAGCGGTTGCTGCTGATC 1633
DB 1644 CCCAGCTGTTTCAGAGACATCTTCTGAGAGGGGCCAGAGGGCTTTTGCAGAGCTG 1703

QY 1634 TCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACTTCCCGGATGCACACCACTT 1693
DB 1704 TGCAGGATCACCAGGGGCTGCTGCTAATGGACACACTTCCGGGATGCCACCACTG 1763

QY 1694 TGCTTGGGACCGGAGTCCGCTCATTCGCACTGAAGCTCGCGGAGAGGCCGCTGCGAAAGC 1753
DB 1764 TACTTGCACATAGAGTGCAGACACGATCTCAAAAAGATTGCAACCTTACGTTGCCACA 1823

QY 1754 TGACTCTGAGCTTTTGTCCGTGGAGGCTGGGCGCGGAGCTAGCATGTGCGATGC 1813
DB 1824 ACTTCAACCACTCTTTCAGCATAGAACTGGGGAGAGCCATTTGACGTGGCCATGC 1883

1814 GTTTCCTCTTTGAGATCCGTTGGGACAGGCTCGACGAGCTGCGGAGCGATGCGGAATG 1873
Db 1884 GCTTCTTTGATGATGCTCCCTGGCGGCTCCAGAGCTCCGGAGCTCATCCCAACA 1943
Qy 1874 TAAACATTCAGATGCTGCTTCGGCGCGCAACACCGTGGATACACCCGTTACCGAGCT 1933
Db 1944 TCCCATTCAGATGCTATCAGGGGGGCCAATGCTGGGGCTACACCACTATCCCTGACA 2003
Qy 1934 CCGTCTCCCGCGCTTTGTTAAAGAACTGCCAGCTCCGCGTGGACATCTTCCGATCT 1993
Db 2004 ACGTGGTCTTCAAGTCTGTGAGGTGGCCAAAGAGAATGCGACGCTTCCCGATCT 2063
Qy 1994 TCGACGGCTTAAAGAGCTTCCAGATGCTCCAGATGCTCCAGCAATCGACGAGTCTCGGAGACCA 2053
Db 2064 TTGACTCCCTTAACTACTGCAAAACATGCTGTGGGCATGGAAGCAG---CTGGCAGTG 2120
Qy 2054 ACACGGGTGAGCGAGTGGCTATGCTTATCTGTGTATCTCTGATCCAAATGAA 2113
Db 2121 CTGGGGTGTGTGAAGCTGCCATCTCTACAGGGGTGAGCTGGCTGACCCAGTCGA 2180
Qy 2114 AGCTCTACACCTGGATTAATCTAAAGATGGACAGAGATCGTCAAGTCTCGCGCTC 2173
Db 2181 CTAATATCTACTGAGATGCTATGAGGCTTAGTGAAGTGTGGAGCGCGCACTC 2240
Qy 2174 ACATTCTGGCCATTAAGATATGCTGCTGCTTCCCGAGCTGGGTAAACCAAGCTG 2233
Db 2241 ACATCTCTGCTATTAAGACATGCGAGCTCTGAAAGCTGCGAGCATGACCATGCTG 2300
Qy 2234 TCACCGCACTGCG---CGGTGAATTCGATCTGCGAGTGCAGTGCACACCCAGCACTG 2290
Db 2301 TCAGCTCCCTCCGGGACCGGTTCCCGACCTCCCACTGCACTTCCATACCCATGACACAT 2360
Qy 2291 CCGGTGGCCAGTGTGGCTACTTACTTGTGCTGAGCTCAAGCTGGTGCAGATGCTGTTGACG 2350
Db 2361 CAGGCTCAGGTGTGGCAGCCATGTTGGCTGTGCACAAAGCTGGGGCTGATGTTGGATG 2420
Qy 2351 GTGCTTCGCACTGCTGCGCACCTTCCAGCCATCCCTGCTGCTGCAATGTTGCTG 2410
Db 2421 TGGCAGTCTGATGCTGCGATGACCTCACAGCCAGCATGGGGGCCCTGTGSCCT 2480
Qy 2411 CATTCGCGCACACCGTCCGATACCGGTTGAGCTTCGAGGCTGTTCTGACCTCGAGC 2470
Db 2481 GTACAAAGGACTCTCTGACACAGAGTACCCCTGAGCGTGTGTTGACTACAGTG 2540
Qy 2471 CGTACTGGGAAGCTGTGGCGGACTGTACCTGCTGCAATTTGA-----GTCTGGACCCAG 2524
Db 2541 AGTATGGGAAGGGCTCGGGGCTGTATGAGCTTTGATGACGCTTACCATGAAGT 2600
Qy 2525 GCCCAACCGGTGCGCTTACCGCCACGAAATCCAGGCGGACAGTTGTCCAACTCGGTG 2584
Db 2601 CTGGCAACTCAGACGTGTATGAGATGAGGATCCAGGGGGCCAGTACCAACACCTACT 2660
Qy 2585 CACAGGCCACCGCACTGGGCTTGTGATCGCTTCGAGCTCATCGAAGACAACTACGAG 2644
Db 2661 TCCAGGCCACAGCATGGGACTTTGGCTCCAAGTTCAAGGAGGTCAAGAAGSCCTATGTG 2720
Qy 2645 CCGTTAATGAGATGTGGGAGCCCAACAAAGGTACCCCATCTCCAAAGGTTGTGGCG 2704
Db 2721 AGGCTAACAGATGTGGGGACCTCATCAAGGTGACACCATCTCTCAAGATTTGGGGG 2780
Qy 2705 ACCTCGCACTCCACCTGTGTGGTGGGTGTAGATTCAGAGCACTTTGTCTGACACCAAC 2764
Db 2781 ATCTGGCCAGTTCAATGTGAGACAGCGGTTGAGCGCGGACAGAGCAAGCTCAGGAC 2840
Qy 2765 AAAAGTACGATCCAGACTCTGTATCGGTTCTCGGCGGAGCTTGGTAAACCTC 2824
Db 2841 AAGACTGCTCTTCCCGCTCTGTGGTGGAGTTCTCGAGGGCTACATTGGCATTCCTCC 2900
Qy 2825 CAGGTGGCTGGCCAGAACCACTGCGCACCCGGGCACTGGAAAGGCC 2869
Db 2901 ATGGGGGTTTCCCTGAACCCCTTCCGTTCTAAGGTGCTAAAGACC 2945

RESULT 6
US-09-880-107-3029
; Sequence 3029, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3029
; LENGTH: 4017
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 S72370
US-09-880-107-3029

Query Match 19.4%; Score 674.2; DB 10; Length 4017;
Best Local Similarity 54.3%; Pred. No. 1.6e-198;
Matches 1523; Conservative 0; Mismatches 1258; Indels 24; Gaps 7;
Qy 86 TCAAAAAGATCTGTGTAGCAAAACCGCGCGGAATCGGGTCCGCTTCGTCGAGCAC 145
Db 192 TCAAAAAGATCATGTGTGGCCAAAGAGTGAGATTCATCCGCTGTTCGGGCTGCA 251
Qy 146 TCAAAACCGGTGCAGCCACGGTAGCTATTTACCCCGGTGAAGATCGGGGATCATTCAC 205
Db 252 CGGAGCTGGGATCCGCAACCGTAGCCATCTCTGAGCAGGACACGGGCCAGATGCA 311
Qy 206 GCTCTTTTCTTCAAGCTGTCCGATTTGTTGTACTGAAGGCTCACAGTCAAGCGTACC 265
Db 312 GGCAGAAAGCAGATGAAGCTATCTCATCGCGCGGCTGCCCCCTGCGAGGCTACC 371
Qy 266 TGGACATCGATGAATATTCGGTGCAGCTAAAAAGTTAAAGCAGATGCTATTTACCCGG 325
Db 372 TGCAATCCAGACATCATCAAGGTGGCCAGGAGAACAAAGTAGATCAGTGCACCCCTG 431
Qy 326 GATATGGCTTCTGTCTGAATATGCCAGCTTCCCGGAGTGCAGCGGAAACGCGCATTA 385
Db 432 GCTACGGTTCTCTGAGCGAGCGGACTTCGCCAGGCTGCCAGATGCGAGGGTCC 491
Qy 386 CTTTATTTGGCCCAACCCCGAGGTTCTTGATCTCACCGGTGATAAGTCTCGTGGGTAA 445
Db 492 GGTATTTGTTCCAGCCCAAGAGTGGTCCGCAAGATGGAGACAAAGTGGAGGCCCGGG 551
Qy 446 CCGCCCGGAAGAGGTGGTCTGCGAGTTTGGCGG---AATCCACCCGAGCAAAAACA 502
Db 552 CCATCGCCATGCTCGGGGTGTCCCGTGGTCCCTGGCAGAGATGCCCCCATCATCTTCA 611
Qy 503 TCGATGACATCTTTAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTAAAGGAGTTG 562
Db 612 TGCATGAGGCCACAGATTCTCCAAACACTACGGCTTCCCCATCATCTTCAAGGGCGCT 671
Qy 563 CCGGTGTTGGCGGAGCGGATGCGCTTGTGTTTCTTTCACCTGATGAGTCCGCAATTTGG 622
Db 672 ATGGGGTGGGGGCTGGCATGAGGGTGTGCACAGCTTACGAGGAGCTGGAGGAGATT 731
Qy 623 CACAGAACATCTCGTGAAGCTGAAGCGGCAATTCGGGACCGTTCGGTATATGTGCAAC 682
Db 732 ACACCGGCGCTACTCAGAGGCTGCGCGCTTGGGAATGGGGCTGTGTTGGAGA 791
Qy 683 GTGCTGTATTAAACCCGAGCAGATGAGTGCAGATCCTTGGCGATCGCACTCGAGAGAG 742

Db 792 AGTTTCATCGAAGCCACCGCACATCGAGTGCAGATCTTGGGGACACAGTATCGGAACA 851
Qy 743 TTGTACACCTTTATGAAGTGAAGTCTGCTCACTGACGCGTGTCAACAAAGTGTTCGAAA 802
Db 852 TCCTGCACTGTACGAGCGAGACTGCTCCATCCAGCGCGGCACAGAAAGTGTTCGAGA 911
Qy 803 TTGCGCCAGCACACATTTGGATCCAGAACTGCGTGTGATCGCATTTGTGCGGATCGAGTAA 862
Db 912 TTGCCCGCGCGCCACCTCGAACCCGAGCTTCGGACTTCGGCTCACCGACACTCTGTGA 971
Qy 863 AGTTTCGCGCTCCATTTGGTTTACCAGGCGCGGGAACCGTGGAAATTTCTTGTCGATGAAA 922
Db 972 AACTCGCTAAACAGGTGGCTTACGGAACCGAGCAACCGTGGAGTTCCTGTGTGACAGC 1031
Qy 923 AGGCAACACAGTTTTCATCGAAATGAACCAACGATATCCAGGTTGAGACACCGTGAATG 982
Db 1032 ACGCAAGACACTACTTTCATCGAGGTCACTCCCGCTGCAAGTGCAGTGCAGTGCAGTGCAG 1091
Qy 983 AAGAAGTCAACGAGGTGAGCTGTGTAAGGCGCAGATCGCGTTCGCTGCTGCTGCAACCT 1042
Db 1092 AGGAGTCAACGAGTGAACCTGTGTCCATGCTCAGATCCACGTGGCTGAAGGCAAGGAGCC 1151
Qy 1043 TGAAGGAATTTGGTCTGACCCCAAGATAAGATCAAGACCCACGCTGCAGCACTGCAGTGCC 1102
Db 1152 TACCGACCTGGGCTGCGGACAGAGACATCCGCATCAACGGGTGTCATCCAGTGCC 1211
Qy 1103 GCATCAACACGGAAGATCCAAACAAAGGCTTCCGCCAGATACCGGAACATATACCCGCT 1162
Db 1212 GGGTCACACCGAGGACCCCGCGCCACGTTCCAGCGGACACCGCGCGCATTTAGGTGT 1271
Qy 1163 ACCGCTCACAGGCGGAGCTGGCGTTGCTCT---TGAGGTGCAGCTCAGCTCGGTGGCG 1219
Db 1272 TCCGAGCGGAGAGGCGATGCGGCTTCCGCTGGATATGCTTCCGCGCTTCCAAAGAGCCG 1331
Qy 1220 AAATCACCGCACATTTGACTTCCATGCTGTGTGAAATGACCTGCGTGGTTCCGACTTTG 1279
Db 1332 TCATCTCGCCCACTACGACTCCCTGCTGTGTCAAGTCAATGTCGCCACGCGCAAGCACCC 1391
Qy 1280 AAATGCTGTGTGTCGACAGCGCGGTGTGGCTGAGTTACCGTGTCTGTGTGTGCA 1339
Db 1392 CCAAGCGCGCACCAAGATGACAGGCGCCCTTGGGAGTTCCGAGTTCGAGGTGTGAAGA 1451
Qy 1340 CCACATTTGGTTTCTTGGTGGTGTGTGCGGGAAGGAGTTCACCTTCCAGCGCATCG 1399
Db 1452 CCAACATGCGCTTCTCGAAGATGTGTCAACACCGAGTTCCTCGCGGCACTGTGG 1511
Qy 1400 CCACCGATTTATCGCGCATCACCCACACCTCTTTCAGGCTCCACCTGCGGATCATGAC 1459
Db 1512 ACACCGAGTTATCGACGAGAACCCAGAGTGTTCAGCTCGGCGCTGCACAGAACCGGG 1571
Qy 1460 AGGACGCATCTGGATTAATTTGACAGATGTCAACCGTGA---CAAGCCTCATGTGTG 1515
Db 1572 CCCAAAGCTGTTCACCTACCTCGGCCATGTATGTGTAAACGGTCCAAACACCCCGATT 1631
Qy 1516 CGTCAAGAGATTTGACGACCAATCGATAGCTGCCCAACATCAAGGATCTGCCACTG 1575
Db 1632 CCGTCAAGGCGAGCCCGAGCCCGACCGACCCGCTGTCTGCTGAGTGCATAGGCGCCG 1691
Qy 1576 CCAG---CGGTTCCGTTGACCGCTGAAGCAGCTTGGCCAGCGCGTTTGTCTCGTATC 1633
Db 1692 CCCCGCTGTGTTTCAGAGACATCTCTGTCGAGAGGGGCTGAGGGCTTTGCTCGAGCTG 1751
Qy 1634 TCCGTGACGAGGACGCACTGGCAGTTACTGATACCATCTTCGCGGATGCACACCAAGTCTT 1693
Db 1752 TGCGGAACCAACCGGGCTGTCTGTGTGAGACACGACCTTCAGGGAACCGCCACAGTAC 1811
Qy 1694 TGCTTGGAACCGAGTCCGCTCATTTGCAATGGAAGCTTCGCGGAGAGCGCGTGCAGAGC 1753
Db 1812 TGCTGGCACTCGTGTGCGCACCCACGATCTCAAAAAGATCGCCCGCTATGTTGCCACA 1871
Qy 1754 TGACTCTGAGCTTTTGTCCGTGAGGCTTGGGCGGCGGACCTACGATGTGCGCATGC 1813
Db 1872 ACTTCAGCAAGCTCTTCAGCATGAGAACTGGGGAGGAGCAACGTTTGAAGCTGCGCATGC 1931

Qy 1814 GTTTCCTCTTTTGGAGATCCGTGGGACAGGCTCGACGAGCTGCGCGGAGCGATGCCGAATG 1873
Db 1932 GCTTCTGTATGATGTCCTTGGCGGCGCTGCGAGGAGCTCCGGGAGCTCATCCCAACA 1991
Qy 1874 TAAACATTCAGATGCTGCTTGGCGGCGCAACACCGTGGGATACACCCGCTACCCAGACT 1933
Db 1992 TCCGTTTCCAGATGCTCTGCGGGGGCCAAATGCTGTGGGTACACCAACTACCCAGACA 2051
Qy 1934 CCGTCTGCCCGCGGTTTGTAAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCGCGATCT 1993
Db 2052 ACGTGGTCTTCAAGTTCTGTGAAGTGCSCAAAGAGATGCGATGATGCTTCCGTGTGT 2111
Qy 1994 TCGACGCGCTTAAAGAGCTCTCCAGATGCGTTCAGCAATCGAGCGAGTCTGTGAGACCA 2053
Db 2112 TTGACTCTCCCTCAACTACTTGTGCCCAACATGCTGTGGGCATGGAGGGG---CAGGAAGT 2168
Qy 2054 ACACCGGGTAGCGAGGTGCTATGGCTTATTTCTGTGTGATCTCTGTATCCAAATGAAA 2113
Db 2169 CCGAGGCGTGTGGAGGCTGCCATCTACACGGGCGAGCTGGCGGACCCAGCGCA 2228
Qy 2114 AGCTCTACACCCCTGGATTAATCTAAAGATGGCAGAGGAGATCGTCAAGTCTTGGCGCTC 2173
Db 2229 CCAAGTACTCACTGCAGTACTCATGGGCTTGGCTGAAGCGCTGGTGGAGCTGGCACCC 2288
Qy 2174 ACATTTGGCCATTAAGGATATGGCTGTCTGTTCCGCCAGCTGCGGTAAACCAAGTTGG 2233
Db 2289 ACATCTGTGCATCAAGGACATGGCGGGCTGTGTGAAGCCACCGGCTGCACCATGTCTGG 2348
Qy 2234 TCACCGCACTGCG---CCGTGAATTCGATCTGCGAGTGCAGCTGCACACCCAGCACACTG 2290
Db 2349 TCAGTCCCTCCGGGACCGTTCGCCGACCTCCCACTGCACATCCACACCCAGCACCGT 2408
Qy 2291 CGGTGGCCAGTGTGGCTTACCTACTTGTCTGAGCTCAAGCTGTGTCAGATGCTTGTGACG 2350
Db 2409 CAGGGCAGCGGTGGCAGCCATGCTGCGCTGTGCCAGGCTGGAGCTGATGTGTGATG 2468
Qy 2351 GTGTTTCGCAACCACTGTCTGGCACCACTCTCCAGCCATCCCTGTCTGCTGCTGTTGCTG 2410
Db 2469 TGGCAGCTGATTCATGCTGGGATGATTTCAAGCCAGCATGGGGGCGCTTGTGSCCT 2528
Qy 2411 CATTCGCGCACACCGTTCGCGATACCGTTTGAAGCTTCGAGGCTGTTTCTGACCTCGAGC 2470
Db 2529 GTACAGAGGAGTTCGCCCTGGACACAGAGTGCCTTGGAGCGCGTGTGTGACTACAGT 2588
Qy 2471 CGTACTGGGAAGCTGTGCGCGGACTGTACCTGCTGCTGCTGCTGGAACCCAGGCCCAA 2530
Db 2589 AGTACTGGGAGGGGCTCGGGGACTGTACGCGGCTTTCGACTGCAACCGCCATGAGT 2648
Qy 2531 CCGG-----TCGCGTCTTACCGCCACGAATCCAGCGGACAGTGTGTCCAACTCGGTG 2584
Db 2649 CTGGCAACTCGGACGTGTATGAAAATGAGATCCAGGGGCGCAGTACACCAACTGCCT 2708
Qy 2585 CACAGGCCACCGCACTGGGCTTGTGATCGCTTCGAGCTCATCGAAGACACTACCGCAG 2644
Db 2709 TCCAGGCCCAACAGCATGGGCTTGGCTCCAAAGTTCAAGGAGGTCAAGAGGCTTATGG 2768
Qy 2645 CCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCAACCATCTCCAAAGTGTGTGCGG 2704
Db 2769 AGGCCAACAGATGCTGGGCGATCTCATCAAGGTGACGCGCTCTCCAAATCGTGGGG 2828
Qy 2705 ACCTCGCACTCCACCTGGTTGGTGCAGGATGATGATCAGCAGACTTTGTGTCAGACCCAC 2764
Db 2829 ACCTGGCCCACTTTATGGTGCAGATGATGAGTGGAGCGGCGAGGCGGAGGCTCAGCGG 2888
Qy 2765 AAAAGTACGACATCCAGACTCTGTATCCGCTTCTGCGCGGAGCTTGGTAAACCTC 2824
Db 2889 AAGAGCTGTCTTTTCCCGCTCCGTGGTGGATTTCTGTCAGGGGTACATCGGTGTCCCC 2948
Qy 2825 CAGTGTGCGGCGCAGAACCACTCGGACCCCGCACTTGGAGGCC 2869
Db 2949 ATGGGGGTTCCTCCCGAACCTTTCTGCTCTTAAGGTACTGAAGGACC 2993

RESULT 7

US-09-815-242-6709

; Sequence 6709, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Karl L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; TITLE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA 011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6709

; LENGTH: 3429

; TYPE: DNA

; ORGANISM: Enterococcus faecalis

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)...(3429)

US-09-815-242-6709

Query Match
Best Local Similarity 14.5%; Score 502.8; DB 10; Length 3429;
Matches 1657; Conservative 0; Mismatches 1672; Indels 51; Gaps 10;

Qy 86 TCAGAAACCGGTGACGACCGGTAGCTATTTACCCCGGTGAAGATCGGGGATCATTCACCC 145

Db 2 TGAAGAAAGGTATTAGTTGCGCAATCGTGGCAATTCGCAATTCGAAATTTTAGAGCTTGTA 61

Qy 146 TCGAAACCGGTGACGACCGGTAGCTATTTACCCCGGTGAAGATCGGGGATCATTCACCC 205

Db 62 CAGAATTAGATATCCGTACAGTAGCGATTTATGCTGCAGAGAGATGAGTATTTCTGTTTATC 121

Qy 206 GCTCTTTTGTCTTCAAGCTGTCGCAATTCGTAAGAGGCTCACCAAGTCAAGGCGTACC 265

Db 122 GTTTCAGAGCAGATCAAGCGTATTAGTTGTTGAAGGAGAAAAACCAATTGAAGCTTATT 181

Qy 266 TCGCATCGATCAATATTCGTGCGAGCTAAAAAGTTTAAAGCAGATGCTATTATACCCGG 325

Db 182 TAGACATTTGAATATTATCCAAATTCGCAAAATTCGGAGCAGATGCGCATTCATCTG 241

Qy 326 GATATGGCTTCTGCTGTAATATCCAGCTTGCCTGCGAGTGGCGGAGGAGGAGGAGGATTA 385

Db 242 GTTATGGCTTTTATCAGAGACTTACGCTTCGCGAACGTTGTGAAGAGAGGAGATCA 301

Qy 386 CTTTATTGCGCCCAACCCAGAGGTTCTTGATCTCACCGGTGATTAAGTCTGTCGGGTAA 445

Db 302 TTTTGTGCGGACCTAAACACATCATTTAGATATTTTGGCGATTAATTAAGCGGAAG 361

Qy 446 CCGCCCGGAGAAAGGCTGCTGTCAGTTTGGCGGAATC---CACCCCGAGCAAAAACA 502

Db 362 AAGCAGCTAGCTGCGGGGATTCGCTTCGATTCAGGCTCAGACGGGCCAGTAGCAACGG 421

Qy 503 TCGATGACATCGTTAAAGCGCTGAAGCGCAGACTTACCCCATCTTTGTAAAGCAGCTTG 562

Db 422 TGGAAAGGTCGTAGCTTTTGGTGAACGATGCGCTTTCCTATCATGATTAAAGCTGCTT 481

Qy 563 CCGTGTGGCGGAGCGGTATGCGCTTGTCTTCTTCACTGATGAGCTCGGAAATGG 622

Db 482 TAGGTGGCGCGGTGCGGGATGCGGTGCGCCAGATGCAAGAGAGCAGGAAAGTT 541

Qy 623 CAACAGAAAGCTCTGTAAGCTGAAGCGGATTCGGGAGCGTTCCGATATATCTCAAC 682

Db 542 ACAGAGAGCCAAAGTGAAGCGAAAGCAGCTTTGGTTCTGACGAGGTTATGTTGAA 601

Qy 683 GTGCTGTGATTAAACCCCGCAGCATTTGAAGTGCAGATCTTTGGCGATGCGACTCGAGAAG 742

Db 602 AGTATATTCTAATCCTAAACATATCGAAGTACAAATTTTAGCGGATCATCATCGGAATG 661

Qy 743 TTGTACACCTTTTATGAAGCTGACTGCTCACTGACGCGCTGTCACCAAAAAGTTTTCGAAA 802

Db 662 TCTTGCAATTTTGAACGCTGATTGTTCCGTTCAACGACGCCCAAAAAGTGTAGAAAG 721

Qy 803 TTGCGCCAGCAGCATTTGATCCAGAACTGCGTATGCGATTCGATTTGTCGGATGCAAGTAA 862

Db 722 TCGACCATGTTGATCAATGAATGAACAAACGAGCAGCATTTGTCGGCTGCTGTCG 781

Qy 863 AGTTCTGCGCTCCATTTGTTTACAGGCGCGGAAACGTCGGAATTTCTTGTTCGATGAAA 922

Db 782 AGTTAATGCGGATGTCGCTACGTAATGCGGAAACGTTGAGTTTGTAGT---AGAAAG 838

Qy 923 AGGGCAACACGTTTTCATCGAAATGAACCCAGTATCCAGTTGAGGACACCGTACTG 982

Db 839 GGGATCAGTTTATTTATTTGAAGTGAATCTCGTTGTTTACGTTAGATCACTACTACAG 898

Qy 983 AAGAACTCACGAGGTGACCTGTTGAAGGCGCAGATGCGCTTGGCTGCTGGTCAACCT 1042

Db 899 AAATGATTACAGATATTGATATTGATCTCTCAATTAACAAATGCAAGGCTGTTGAT 958

Qy 1043 TGAAGGAA-----TTGGGTCTGACCCAAAGATAAGATCAAGACCCACCGTCAGCAGCTGC 1096

Db 959 TGCATAAAGATATGCAATTTGCAAAACAAACGAATTTGACATTTAAAGGCGCGGTATTC 1018

Qy 1097 AGTGGCGGATCACACGGAAGATCCAAACAAACGCGCTTCGCGCCAGATACCGGAACATCA 1156

Db 1019 AATGTCGGATTACGACTGAAGATCCCTGAAACAGTTTATGCAAGATACAGGGAATTCG 1078

Qy 1157 CCGCTACCGCTCACCGCGGAGCTGCGCTTCGTTTGAACGGTGA---GTCAGCTCG 1213

Db 1079 ATAGTATCGTTTACAGAGTGGTTGTTGTTGCGTTTGAATGATGAGAAATGCTTATTCG 1138

Qy 1214 GTGGGAAATCACCGCACCTTTGATCCATGCTGCTGAAATGACCTGCGGTGCTTCG 1273

Db 1139 GCTACGCTGACGCTTACTTTGATTTCTTTATGTTTAAAGTCTGTACGATGCTTCT 1198

Qy 1274 ACTTTGAAATCGTGTGCTGTCAGCGCGCTTGGCTGAGTTTACCGTGTCTGCTG 1333

Db 1199 CTTTGAACAGGCAATCAAGTAAATGCAACGCTGCTTAAAGAAATTCGGAATTCGTTGCG 1258

Qy 1334 TTGCAACCAACATTTGTTTCTTGGCTGCTGCGGGAAGAGACTTTCATTTCAAAGC 1393

Db 1259 TGAACCAAAATTTTCGTTTTCACAAATGTTGTGAGCTATCCAGCTTTCATCTGCGG 1318

Qy 1394 GCATGCCACCGGATTTATCGGCGATCACCCACACTCTTTCAGGCTCCACCTCGCGGATG 1453

Db 1319 AAGCCAAAACAACTTTTATGATAATACACCTGAAATTTTGAATTTCTCTGATGCGCG 1378

Qy 1454 ATGAGCAGGAGCGCATCTCGGATTTCTTGGCAGATGTCACCGTGAACCAAGCTCATGTTG 1513

Db 1379 ATCGTGGCAATAAAGCATGAATCATTTGGAGAGTAACCGTCAATGGTTTCCCTGGCA 1438

Qy 1514 TGGTTCCAAAGG-----ATGTTGACGACCAATCGATAAGCTGCCCAACATCAAGGATC 1567

D	b	1439	TTGAACGAAACGGAGAAAAATATTTTGAAGCCCGCGCTGCCGACAGATATTGAAGTTC	1499
Q	y	1568	TGCCACTGCCACGCGGTTCCCGTGACCGCTGAAGCAGCTTGCCCGACCGCGTTTGCTC	1627
D	b	1499	CAGAAAAGTGATCACAGCTAAAAAATATTTAGACGCTCAAGGACCAACCGCGGTATCG	1558
Q	y	1628	GTGATCTCCGTGAGCAGGACGCACCTGGCAGTTACTGATACACACTTTCGCGATGCAACC	1687
D	b	1559	ATTGGGTTAAAGATCAAGAAAGTGTGTAAATGACAGATACCAAGTTCGGTATCGCACC	1618
Q	y	1688	AGTCTTTGCTTGGACCCGAGTCGCTCATTTGCACTGAAGCTCGCGCAGAGCCGCTCG	1747
D	b	1619	AAAGTTTATTAGCTACTCTGTGTGAGAACACAGATTTTAAAGCAATGCTCGACTGA	1678
Q	y	1748	CAAAAGCTGACTCTCTGAGCTTTTCTCGTGGAGGCGCTTGGGCGCGCGACCTACGATG	1807
D	b	1679	ATGTAGCCCTCCCTGAGCTGTTTCTAGTGAATGTGGGTGCGCTACGTTCCGATTTG	1738
Q	y	1808	CGATCGGTTTCTCTTTTGAGGATCCCGTGGGACAGGCTCGACGAGCTGCGGAGCGATG	1867
D	b	1739	CTTATCGCTTCTTAAACCGAAGACCCATGCGCAACGTTTAAAGAAAAATTCGTCA	1798
Q	y	1868	CGAATGTAACATTTCAAGTCTGCTTTCGCGCGCGCAACACCGTGGGATACACCCCGT	1927
D	b	1799	CAAAACATCTTTTACAAATCTCTGTCAGAGGATCTAATGCGGTAGGCTATCAAA	1858
Q	y	1928	CAGACTCCGCTGCGCGCGCTTTGTTTAAAGAAAGCTGCCAGCTTCCGGCTGGACAT	1987
D	b	1859	CTGATATGTTATTGAAGAATTCATTTAAGATCTGCTCGCCAGGAGTCAATGTTTCC	1918
Q	y	1988	GCATCTTCGACGGCTTTAACGACGCTCTCCAGATGGCTCCAGCAATCGACGAGTCT	2047
D	b	1919	GAATTTTGTAGCTTAAACTGGAATCTCTCAAAATGGAATAAGTATTAAGTCGTCG	1978
Q	y	2048	AGACCAACACGGGTAGCCGAGGTGGCTATGGCTTATTTCTGGTGATCTCTCTGATC	2107
D	b	1979	ATACCGGAA--AAATGCGGAAGCAGCAATTTGTTATCTGGGGACATCAATGATCC	2035
Q	y	2108	ATGAAAAGCTCTACACCTCGGATTACTCTAAAGATGGCAGAGGAGATCGTCAAGTCT	2167
D	b	2036	CCCGAGCAAAATATATGTTCAATATTAATCTTGATATGGCTAAGATTTGAAAAATT	2095
Q	y	2168	GGCTCTCATTTCTGGCCATTAAGGATATGGCTGGTCTGCTTGCCCGCAGCTGCGGTAA	2227
D	b	2096	GGGCACAAATCAATTGGCATTTAAGATATGGCTGGCTTATTGAAAACCAAGCTCTTAT	2155
Q	y	2228	AGCTGGTCAACGCACTCGCGCGGTGAATTCGATCTGCCAGTGCACGTGCACACCA	2287
D	b	2156	GTTTTAAATTAGTGAATTTAAAGGACGCCAGGATTTACCAATTCATCTCCATACT	2215
Q	y	2288	CTGCGGTGGCAGTTGGCTTACCTACTTTGTCGAGCTCAAGCTGGTGCAGATGCTGTG	2347
D	b	2216	CTAGTGCAATGGGATCATCTTATTACAGAGCTACTTAAGCAGGCGTTGATATTGTTG	2275
Q	y	2348	ACGCTGCTCCGCAACCATCTGTGCGCACCACTCCAGCCCATCCCTGCTGCGCATTTG	2407
D	b	2276	ACGTCCGAATGAGTGGCATGAGTGGCGCAACTAGCCAGCAAGTATGAACAGTCTAT	2335
Q	y	2408	CTGCATTCGCGCACACCCGTCGCGGATACCGGTTTGAGGCTCGAGGCTGTTCTGAC	2467
D	b	2336	ATGCTTTAGTCAATGTTGTAACGCGCAACCTATTTAACTTGAATATGCAAGAAAA	2395
Q	y	2468	AGCGTACTGGGAAGCTGTGCGCGGACTGTACTCTGCAATTTGAGTCTGGAACCCAG	2527
D	b	2396	ATCATTTATGGGAAGATGTGCGCATGTATTATCAACCAATTTGAAAAATGGGTTAA	2455
Q	y	2528	CAACCGGTCCGCTTACCGCCACGAATCCAGGCGGAGTAGTTGTGTCAACTCGCTGC	2587
D	b	2456	CGCAAAACAGAAAGTCTATATGCAATGAAATGCGCTGGTGGTCAATACTCTAATCT	2515
Q	y	2588	AGGCCACCGCATCTGGGCTTGTGATCTGCTTCGAGCTCATCGAAGACAACTACGCG	2647
D	b	2516	AAGCAAAAGCGGTAGGCTTAGGGCACCGTTGGGATGAAATCAAAAAATGTATCACAC	2575

QY	2648	TTAATGAGATGCTGGGACGCCCAAACAAGGTCAACCCATCTCTCAAGTGTGTGGCGACC	2707
DB	2576	TCGAATTTGATGTTTGGTGATATTGTAAAAAGTACGCCATCTTTCAAAAGTCGTGGGAGACA	2635
QY	2708	TCGCATCTCCACCTGGTGTGGTGTAGATCCACGACAGCTTTGCTGCAGACCCACAAA	2767
DB	2636	TGGCACTCTTCAATGTTCAAAATCATCTCAGACAGACAGATGTTTATGCGGTGGTGAAG	2695
QY	2768	AGTAGACATCCAGACTCTGTCAATGCGCTTCTCGCGCGAGCTTGGTAAACCTCCAG	2827
DB	2696	AACTAAGTTTTCCTGAATCAGTTGTGACTTTTTTCCAAAGTGATTTAGGTCAACCAAGTTG	2755
QY	2828	GTGSGCTGGCGAGAACCACTGSCGCAACCCGCGCATGGAAGCCGCTCCGAAGGCAAGGCAC	2887
DB	2756	GTGGGTTCCTCCGAAGAGCTCCAAACGAATTAATTTCTAAAGGTCGACGAGCTTCACAGAGC	2815
QY	2888	CTCTGACGGAAGTTCTCTGAGGAAGAGCA---GGCGCACCTCGACGCTGATGATTCCAAG	2943
DB	2816	GGCCAGGTGACTTAGCAGACACTGTTGATTTTTGCGNAAAGTACAAGAAGATTAAGCCGAAA	2875
QY	2944	GAAAGTTCGCAACAGCCTCAACCGCTGCTGTTC-----CGAAGGCCAA	2986
DB	2876	AAATTTGGGTATCAACCTAAATTTGGAAGAAGTTTAAAGTTATTTAATGTATCCACAAGTAT	2935
QY	2987	CCGAAGAGTTTCTTCGAGCAACCGTCGCGCTTCGGCAACACCTCTGCGCTGGATGATCGTG	3046
DB	2936	TTTTTAGAGTATCGTCAAAATAACGAAACCTTTGGAGATATAACTTTATTAGATACGCCAA	2995
QY	3047	AAATTTCTTACGAGCTGTCGAGGCGCGGAGACTTTGATCGCCCTGCCAGATGTGCGCA	3106
DB	2996	CATTCTTTAAATGATTCGTCAAGGA--GAGACATTGGAAGTTCAAATTTGAACGAGGAA	3052
QY	3107	CCCACTGCTTGTTCCTCGATGCGATCTCTGAGCCAGACAGATAAGGGTATGCGCAATG	3166
DB	3053	AAACGTTAATTAATCGTTTAGATGAATTTGGCGAACCTGATATTGACGGAACCGTGCTCT	3112
QY	3167	TTGTGGCCAAAGCTCAAGSCCAGATCCGCGCAATGCGTGTGCTGAGCCGCTCCGTTTGAGT	3226
DB	3113	TGTTCTTCAATTTGAAATGGGCAACGTCGTGAGGTTTTAGTTAAAGATGCCCTCAATTAAGT	3172
QY	3227	CTGTCAACCGCAACCCAGAAAAGGAGATTCCTCCAACAAAGGGCCATGTTGTGCAACCAT	3286
DB	3173	CTGCTGTTCAAGTAAAAAATAAAGGCGGAACCAACTAAATAAGAACAAATTCGGCGCAACGA	3232
QY	3287	TGCTCTGG---TGTTGTCACTGTGACTGTGCTGAAGGTGATGAGGTCAAGGCTGGAGATG	3343
DB	3233	TGTCGTGGTTCGTTCTCAACGATTTTGGTCAAAACGTGGCGATAAAGTAGAAAAAGGTGACG	3292
QY	3344	CAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACATCACTGCTTCTGTTGACGGCA	3403
DB	3293	CGTTGCTGATTACAGAGCCATGAATAAGAAACGACTATCGAAGCACGTTTTTGGCGGCA	3352
QY	3404	AGATTGAACGCGTGTGGTT	3423
DB	3353	CGTTTGATCACATTTATGTT	3372

```

RESULT 8
US-08-781-986A-324/c
; Sequence 324, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 324:
SEQUENCE CHARACTERISTICS:
LENGTH: 5030 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-324

Query Match 14.1%; Score 488.4; DB 7; Length 5030;
Best Local Similarity 48.5%; Pred. No. 1e-140;
Matches 1665; Conservative 0; Mismatches 1711; Indels 60; Gaps 9;
QY 84 ATTCAAAAGATCTTGGTAGCAACCCGCGCAATCGCGTCCGCTTTCGCGTGCAGC 143
DB 3846 AATAAAAAAGTTACTTGTGTAAACCGTGGAGAAATTCGAATTCGATATTCAGAGCGGC 3787
QY 144 ACTCGAACCCGCTGCAGCAGCGTAGCTATTTACCCCGTGAAGATCGGGGATCATTTCCA 203
DB 3786 GGCAGATTAGACATCAGCACAGTTGCAATTTATTGCAATGAAGACAAAGATTCAATACA 3727
QY 204 CGCTCTTTTCTTGAAGCTGTCGCGATTTGGTACTTAAGGCTCACCAGTCAAGGGTA 263
DB 3726 TAGATATAAGCAGATGAATCTATTAGTTGGAAGTATTAGTCTCTGCTGAAAGTTA 3667
QY 264 CCGTGCATCGATGAATATTCGCTGAGCTAAAGATTAAGCAGATGCTATTTTACCC 323
DB 3666 TTAAATATTAGCGTATCATTTAGTAGAGAAACACAGCGATGTGATGCGATTCATCC 3607
QY 324 GGGATATGCTTCTGTCTGAATATCCAGCTTCCCGCGAGTCCGCGGAAACGGCAT 383
DB 3606 TGGCTATGGAATTTTAAGTGAATATGAACAATTTGCGCTGCTGTCGAGAGAGGAT 3547
QY 384 TACTTTTATGGCCCAACCCAGAGGTTCTTGATCTCACCGGTGATAGTCTCGTGGGT 443
DB 3546 TAAATTTATTTGCTCTCATTTTAGAACATTTAGATATGTTGGAGATAAAGTTAAGCTCG 3487
QY 444 AACCCTCCGGAAGAGGCTGCTCCAGTTTGGCGGAATCCACCCCGAGCAAAAACAT 503
DB 3486 TACAACCGGTATCAAGGCGAGATTTACCAGTTATTCTCGTACAGACGGTCCAATTAATC 3427
QY 504 CGATCAGATCGTTAAAGCGCTGGAAGGCGAGAC---TTACCCCATCTTTTGAAGGCGAGT 560
DB 3426 ATATGAATTAGCAAAAGAAATTTGCAGAAAGCTGGTTTCCCGCTTAATGATTAAGCCAC 3367
QY 561 TCCCGTGTGGCGGACGCGGTATCGCTTTTCTTCACTGATAGCTCCGCAAAAT 620
DB 3366 AAGTGTGGCGCGGTAAAGGTATGGAATCGTTCTGTAAGAAAGTGAATTAGAAGATGC 3307
QY 621 GGCACAGAGGATCTCGTAGCTGAAGCGCATTCGCGACGCTTCGGTATATGTCGA 680
DB 3306 TTTCCATAGAGCAAAATCAGAGCTGAATAATCAATTTGGTAATAGTGAAGTTTACATAGA 3247
QY 681 AGTGTGTGATTAACCCCGACGATTCAGTGCAGATCTTGGCGATCGCACTGGAGA 740

DB 3246 AAGATACATTGTATATCCAAAGCATAATTGAAGTACAGTCACTAGGTGACGAACATGGA 3187
QY 741 AGTTCTACACCTTTATGAACGCTGCTCTCACTGAGCGTCTGTCCACCAAAAAGTTGTGGA 800
DB 3186 TATCGTACACTTATTTGAACGCTGATTGTTTCAAGTCAACGCTGCTCATCAAAAAGTTGTAGA 3127
QY 801 RATTGGCGCAGCACAGCATTTTGGATCCAGAACTGCGTGCATCGCATTTTGTGGGATGCACT 860
DB 3126 AGTTGACATCATGTTGGATATCATCAACATTAGTCAACGTATTTGTGATGCTGCAAT 3067
QY 861 AAAGTTCTGCGCTCCATTGGTTTACAGGGCGCGGAAACCTGGGAATCTTGTGTCGATGA 920
DB 3066 TCAATTTGATGGAAATATTAATATGTCAATGCAAGTACTGTTGAATTTCTAGTATCTG- 3008
QY 921 AAAGGCAACCACTGTTTCAATCGAAATGAACCCAGTATCCAGTTTGAAGCACACCGTGAAC 980
DB 3007 --GTGACGAATTTCTTTTATAGAAATTAACCTCGTGTACAAGTAGAGCATACAAATTAC 2950
QY 981 TGAAGAAAGTCAACGAGGTGGACCTGCTGAAGCGCAGATGCGCTTTGGTGTGCTGTCGAAC 1040
DB 2949 AGAGATGGTAAACAGGAATTTGATTTTAAAGACACAAATTTTAGTTGACGACGGTCCGA 2890
QY 1041 CTTGAAGGAATTTGG-----GTCTGACCCCAAGATTAAGATCAAGACCCACGGTGCAGC 1091
DB 2889 TTTATTTGGTGAAGAGATTTAATATGTCGCAACCAAAAAGATATTACAACATTAGGCTATGC 2830
QY 1092 ACTGAGTCCCGCATCACCACGGAAGATCCAAACAAACGGCTTCCGCCAGATACCGGAAC 1151
DB 2829 CATCCAATGCTGATATACACAGAGATCCGTTAAATGATTTTCAATGCCGATCTGGAAC 2770
QY 1152 TATCACCGGTACCGTCAACAGCGGAGCTGGCTTCTGCTTTCGCTGACGTTGACGTCAGCTCAGCT 1211
DB 2769 AATCATGCTTATCGTTTCAAGCGGGCTTTGCTGATACGCTTTGATCTGGAGATGCTTT 2710
QY 1212 C---GGTGGCGAAATCAACGCAACATTTGACTCCATGCTGGTGGTGAATACCTGCGGTGG 1268
DB 2709 CCAAGTGTGAGATATCACTTATTATGATTCATTACTCTGTAATAATATATCTACACACG 2650
QY 1269 TTCGCACTTTGAACTGCTGTTGCTGTCGACAGCGCGGTTGGCTGAGTTCCACCGTGC 1328
DB 2649 GATATCATTTAAACAGCAGAGAAAGAAATGATGCTCATTTACGAGAAATGCTATTCG 2590
QY 1329 TGGTGTGCAACCAACATTTGTTTCTTGGCTGCTGCTGTCGCGGAGAGAGACTTCACTTC 1388
DB 2589 TGGTGTAAACACTAATATTTCCATTTCTTAATTAATGTAATGAAGAATAAAGTTCAACAG 2530
QY 1389 CAAGCGCATCGCCACCGGATTTATCGCGATCACCCACACCTCTTTCAGGCTCCACCTGC 1448
DB 2529 TGGTGTATTACACAACTAAATTTATTGAAGAAACACCAAGAACTTTTCGACATTCAGCCGTC 2470
QY 1449 GGATGATGACGAGGACGATCTCGATTTACTTGGCAGATGTCACCGTGAAACAAAGCTCA 1508
DB 2469 TCTAGATAGAGTACTAAACATATTAGTAAATAGTAACAATTAATGTTTTC 2410
QY 1509 TGGTGT-----CGGTCCAAAGAGTGTGCGACCAATCGATAAGCTGCCCA 1556
DB 2409 AATGTTGAGAAAGTCCGAAACAGCATATGATTAATGATCANATCCCACTGATCTTC 2350
QY 1557 CATCAAGGATCTGCCACTGCCACGCGGTTCCTGTCAGCGCTGAGCAGCTGGCCGAGC 1616
DB 2349 AAGTAAATCGCTTCATTT---AGTGTACGAAACAAATTTGCTGATGAGTAGTCAAA 2293
QY 1617 CGGTTTGTCTGATCTCGTGAGCAGACGCACTGGCAGTTTACTGATACCACTTCG 1676
DB 2292 AGGTGAGTGAATGGGTTTAAAGAGCAGGATGATGTCTTACTAACAGATACCAACCTTTAG 2233
QY 1677 CGATGACACACAGTCTTTTGTCTGCGACCCGAGTCCGCTCATTCGCACTGAAGCTCGGC 1736
DB 2232 AGATGCACACCAATCATATTAGCTACACGAGTTAGAACTTAAGGATATGATTAATTCGC 2173
QY 1737 AGAGCGCTCGCAAGAGCTGACTCTGAGCTTTTGTCCGTGGAGCCCTGGGGCGCGCGAC 1796
DB 2172 ATCCAAACACGCGACGTAATTTAAAGATGTTTCTCACTAGAAATGTGGGCGGTGCTAC 2113

Qy 1797 CTAGATGTGGGATGCGTTCTTCTTTGAGGATCCGTFGGGACAGGCTCGACGAGCTGG 1856
Db |||||
2112 ATTTGATGTGCATATATTTCTTGAAGAAAAACCCATGGGAACGACTTGAACGCTACG 2053
Qy 1857 CGAGCGATCGGGAATTAACATTTAGATGCTGCTTCGCGCGCGACACCGTGGGATA 1916
Db |||||
2052 TAAGCTATTCGAATGATTTATTTCCAAATGTTGTTAGCTCTTCAACGCGAGTTGGTTA 1993
Qy 1917 CACCCGTACCCAGACTCCGCTGCGCGCGCTTTGTTAAGGAAGCTGCCAGCTCCGGCGT 1976
Db |||||
1992 TAAAACTATCTGATATGTTTATTAATAATTCGTACAAGAAATGCTAAAGCAGGAT 1933
Qy 1977 GGACATCTTCGCGATCTTCGAGCGGCTTAAACGAGTCTCCAGATGGCTCCAGCAATCGA 2036
Db |||||
1932 AGATGCTCTTTAGAAATTTTCGATTCATTTAAACCTGGGTAGATCAAAATGAAGTTGCCAATGA 1873
Qy 2037 CGCAGTCTCGGAGACCAACACCGCGGTAGCGAGGTGCTATGCTTATTCGTGTGATCT 2096
Db |||||
1872 AGCAGTACAAGAGCGGCAAAATCTCAGAAGGTACTATTTGTTATACAGGTGACATTTT 1813
Qy 2097 CTCTGATCCAAATGAAAGCTCTACACCTCGGATTACTACCTAAAGATGGCAGAGGAGAT 2156
Db |||||
1812 AAATCTGAGCGATCAAACTTTATCTTTAGAGTATTTATGCTCAACTAGCTAAAGATT 1753
Qy 2157 GGTCAAGTCTGGCGCTCACATCTTGCGCATTTAAGATATGGCTGGTCTGCTTCCGCCAGC 2216
Db |||||
1752 AGAACGTGAAGGTTTCCATATTTTAGCGATTAAGATATGCGAGGCTTATTTAAACCTAA 1693
Qy 2217 TGGGTAACCAAGCTGGTCAACCGCACTGCGCGGTGAATTCGATCTGCGAGTGCAGTGCA 2276
Db |||||
1692 AGCGCTTACGAATGATTTGGTGAGTTAAATACAGTGTAGATTTACCAATTCATCTTCA 1633
Qy 2277 CACCCAGACACTCGGGGTGGCAGTTGGCTACCTACTTTGCTCGAGCTCAAGCTGGTGC 2336
Db |||||
1632 CACTCATGATACAGTGTAAATGTTTATTAACATACAAACAGCAATAGATGCTGGTGT 1573
Qy 2337 AGATGCTGTGACGGTGTCCGACACATGCTTGCGACACACTCCAGCGCATCCCTGTC 2396
Db |||||
1572 CGATATCATTTGATGACTGCTGTGCTCAATGAGTGGTTTAAACAGTCAAGCGGCA 1513
Qy 2397 TGCCATTTGCTGCTCATTCGCGACACCGTGGCGGATACCGGTTTGAGCCTCGAGGCTGT 2456
Db |||||
1512 TTCTGTTATTTATGATTAATGCTTCCACGCCACCTTAGAACTGATATTTGAAGGTAT 1453
Qy 2457 TTCTGACCTCGAGCGGTACTGGGAAGCTGTGCGGGACTGTACCTGCGCATTTGAGTCTGG 2516
Db |||||
1452 GGAGTCACTTAGTCAATTTGTTGTTCACTGTAGCTACTTATTTACAGCTTTGAAAGTGA 1393
Qy 2517 AACCCAGGCCCAACCGGTGCGGTCTACCGCCACGAAATCCAGCGGACAGTTGTCCAA 2576
Db |||||
1392 TATCAAAATCACCGAATACTGAAATTTATCAACATGAATGCTGCTGGACAGTATTCGAA 1333
Qy 2577 CCTGCGTGACAGCCACCGCACTGGGCTTGTGATGCTGTGAGTCTGAGTCAATCGAGACAA 2636
Db |||||
1332 TTTAAGTCAACAGCTAAAAGTTTAGTTAGGCGAAGATTTGATGAAGTCAAAAGATAT 1273
Qy 2637 CTAGCGAGCGGTTAATGAGATGCTGGGACGCCCAACCAAGTCAACCCATCTCCAGGT 2696
Db |||||
1272 GTATCGCAGAGTGAATTTCTTATTTGGTGATATCGTAAAGTAAACATCTGCTTAAAGT 1213
Qy 2697 TGTGGGACCTCGCATCTCCACTGCTGGTGGTGGGGTGTAGATCCAGCAGACTTTGCTGC 2756
Db |||||
1212 AGTTGGTGATATGCACTTTATATATGTTACAAAATGATCTTTGATGAACAAATCCGTTGATTAC 1153
Qy 2757 AGACCCACAAAGTACACATCCAGACTCTGTCTGCTGCTGCTGCGCGGAGCTGG 2816
Db |||||
1152 AGATGGCTATAAATTTAGATTTCCAGAAATCAGTAGTCTGCTTCTTCAAAAGGTGAATAGG 1093
Qy 2817 TAACCTTCCAGGTGGTGGCCAGAACCACTGCGCACCGCGCTACTGGAAGCGGCTCCGA 2876
Db |||||
1092 ACACCTGTAATGGTTTTTAATAAGATTTTACAGCGGTTATTTTAAAGGCCAAGNAGC 1033

Qy 2877 AGGAGGCACTCTCTGACGGAAGTTCTCTGAGGAAGAGCAGCGCACCTCGACGCTGATGA 2936
Db |||||
1032 ACTAAACAGCTCGTCCAGTGAATATCTAGAGCCAGTTGATTTTGAAGAAAGTCCTGAGTT 973
Qy 2937 TTCCAGGAACTGCGCAACAGCC-----TCAACCGCTGCTGTT 2975
Db |||||
972 GCTTGAAGAAGAGCAACAGGTCCTGTTACGGAGCAAGATATTTATTAGTTATGTTACTATA 913
Qy 2976 CCCGAAGCCAAACCAAGAGTTCTCTCGAGCACCGTTCGCGCTTCGGCAACCACTCTCGCT 3035
Db |||||
912 TCCAAAGTATATGAACAATATATTCAAACCTAGAAATCAATACGAAACCTTATCGTTACT 853
Qy 3036 GATGATGCTGAATTTCTTACCGGACTGTGAGGGCCGAGACTTTTGATTCGCGCTGCC 3095
Db |||||
852 TGATGCGCTTACATTTCTTTTGGAAATGCGTAATG---GTGAACACAGTAGAATTCGAAAT 796
Qy 3096 AGATGTCGCAACCCACTGCTGCTTCGCTGGATGCGATCTCTGAGCCAGACAGTAAAGG 3155
Db |||||
795 CGATAAGGTAACGATTAATTATTAACCTAGAAACGATTTAGTGAACCCAGATGAAATGG 736
Qy 3156 TATGCGCAATGTTGTGGCCAAACGTCACCGCCAGATCCGCCCAGTGGTGTGGTGACCG 3215
Db |||||
735 TAATGAGCAATTTACTATGCGATGATGCTCAAGCGAGCTGTTTACATTAAGATGA 676
Qy 3216 CTCGTTGAGTCTGTCACCGCAACCGCAGAAAGGAGGATTCCTCCAAAGGGCCATGT 3275
Db |||||
675 AAATGTCATACAAATCGCAACGTTAAGCCAAAGCAGATAAGAGTAATCCAAGTCAAT 616
Qy 3276 TGCTGCAACATTCGCTGGTGTGCTCACT---GTGACTGTTGCTGAAGGTGATGAGTCAA 3332
Db |||||
615 CGGTGCTCAAAATCCAGGTTCACTGAAGTCAAGGTTAGTGTAGGTGAGTGAACCTGTGAA 556
Qy 3333 GGCTGGAGATGCACTGCGCAATCATCGAGGCTATGAAGATGGAAGCAACATCACTGCTTC 3392
Db |||||
555 AGCTAATCAGCCGCTGCTTAATTAAGCTATGAAGTGAAGCAACCAATTCAGGACC 496
Qy 3393 TGTGACGCGCAAGATTGAACGCTTGTGTTCTGCTGCAACGAAAGTGAAGGTGGCGA 3452
Db |||||
495 ATTTGACGGTGTGTTAAACCAAGTAACTGTAATAATGTTGATGACACAATAGCAGCGCA 436
Qy 3453 CTTGATGCTGCTGCTT 3468
Db |||||
435 TTTATTAAATCGAAAT 420

RESULT 9

US-09-070-927A-89/c
; Sequence 89, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
; Patrick J. Dillon
; Steven Barash
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 982
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,927A
; FILING DATE: 04-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/046,655

FILING DATE: 1997-05-16
APPLICATION NUMBER: 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB369
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 20072 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 89:

US-09-070-927A-89

Query Match 14.1%; Score 488.2; DB 10; Length 20072;
Best Local Similarity 50.3%; Pred. No. 2,5e-140;
Matches 1398; Conservative 0; Mismatches 1358; Indels 25; Gaps 7;
QY 86 TCRAAAGATCTTGGTAGCAAAACCGCGCGAATTCGCGTCCGTTCCGTTGACGAC 145
DB 5936 TGAATAAGGTATTAGTTGCCAATCTGCGCGAATTCGATTCCGAATTTTAGAGCTTGTA 5877
QY 146 TCGAAGCGGTGACGACCGTAGCTATTACCCCGTGAAGATCGGAGTCAATCCACC 205
DB 5876 CAGAAATAGATATCCGTACAGTAGCAATTTATGTCGAGAAGATAGTATTCTGTTCATC 5817
QY 206 GCTCTTTTGTCTGAAAGCTGTCCGCAATGGTACTGAAGCTCACAGTCAAGCGTACC 265
DB 5816 GTTTCAAGCAGATGAAGCGTATTAGTTGGTAAAGGGGAAAAACCAATTGAAGCCTATT 5757
QY 266 TGGACATCGATGAATATTACGTTGACATAAAAGTTAAAGCAGATGCTATTATCCCGG 325
DB 5756 TAGACATTGAAATATTATCCAAATTGCCAAAAATCGGAGCAGATGCCATTATCCTG 5697
QY 326 GATATGGCTTCTCTGTGAATAATCCGAGCTTGCCCGAGTGGCGGAAAAACGGATTA 385
DB 5696 GTTATGGCTTTTATCAGAGAACTTAC-GTTGCGCGAAAGCTTGTGAAGAGAGGAATCA 5638
QY 386 CTTTATTGGCCCAACCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGTGGGTAA 445
DB 5637 TTTTGTGCGACCTAAACACATCATTTAGATATTTTGGCGATAAAATTAAAGCGAAG 5578
QY 446 CCGCGCGAAGAGGCTGTTCTGCCAGTTTGGCGGAATCCACC---CCGAGCAAAACA 502
DB 5577 AAGCAGCTGATGCTGCGGAGTTGCTTCGATTCCAGCTCAGACGGGCCAGTAGCAACGG 5518
QY 503 TCGATGATCGTTAAAGCGCTGAAGCCAGACTTACCCATCTTTGTAAAGCGAGTTG 562
DB 5517 TGAAGAGGTTGTAGCTTTTGGTGAACACATATGCGTTTCTATCATGATTAAAGCTGCTT 5458
QY 563 CCGGTGTGGCGGACGCGGTATGCGTTTGTCTTCCACTGATAGCTCCGCAAAATGG 622
DB 5457 TAGTGTGGCGCGTGGCGGATGCGGTTGCCACAGATGTCAAAGAGACACAGAGAGTT 5398
QY 623 CAACAGAGCATCTCGTGAAGCTGAAGCGCATTCGCGCAGCGTTGCGTATATGTCGAAC 682
DB 5397 ACGAAGAGCAAAAAGTCAAGCAAGACGAGCCTTTGCTTCTGACGAGTTTATGTTCAA 5338
QY 683 GTGCTGTGATTAAACCCAGCAGATCAAGTGAAGTCAAGTCTTGGCGATCGCATCGAGAAG 742
DB 5337 AGTATATTTCTAACTCTAAACATATCAAGTACAAATTTTAGCGCATCATATGGGAACG 5278
QY 743 TTGTACACCTTTATGAAGCTGACTCTCACTGAGCGCTGTCACCAAAAAGTTGTCGAA 802
DB 5277 TCTTGCAATTTATTTGAACGTGATTGTTCCGTTCAACGGCCGACCAAAAGTGTAGAAG 5218

QY 803 TTGGCCAGCACAGCATTTGGATCCAGAACTGGGTGATCGCATTTGTGCGGATGCAATAA 862
DB 5217 TCGCACCATGTGTATCAATGAATGAAGAAACAACGAGCAGCAATTTGTTCGGCTGCTGTCG 5158
QY 863 AGTTCTCGCGCTCCATTGGTTTACCAGGGCGCGGAACCGTGAATTTCTTGTCGATGAAA 922
DB 5157 AGTTAATGGCACATGTCGGCTACGTGAATGCGGAACGGTTGAGTTTTAGT---AGAAG 5101
QY 923 AGGCAACAGGTTTTCATCGAATGAACCAAGTATCCAGGTTGAGCACCGTGAATG 982
DB 5100 GGGATCAGTTTTTATTTATTTAGTGAATCTCTGTTTCAAGTAGAACATCTATCAAG 5041
QY 983 AAGAAGTCCCGAGGTGACCTGGTGAAGCGGAGATGCGCTTGGCTGCTGGTGAACCT 1042
DB 5040 AAATGATTTACAGATATTGATATTGTGATCTCTCAATTTACAAATTCGCAAGGGCTTGA 4981
QY 1043 TGAAGGAA-----TTGGGTCTGACCCCAAGATAAGATCAAGACCCACGGTGCAGCACTGC 1096
DB 4980 TGCATAAGATATGCAATTTGCCAAAAACAACGAATTGACATTTAAAAAGGCGCGCTATTC 4921
QY 1097 AGTCCCGCATCACCAACGAAAGATCCAAACACCGCTTCCGCCAGATACCGGAACATCA 1156
DB 4920 AATGTCGATTAACGACTGAAGATCCCTGAAACAGTTTATGCCAGACACACAGGAAAAATCG 4861
QY 1157 CCGGTACCGCTCACACGCGGAGCTGCGTTGCTCT---TGACGGTGCAGCTCAGCTCG 1213
DB 4860 ATACGTATCGTTTACACAGGTGTTTGTGTGCGTTTGTAGATGTAGGGAATGCTTTATTCG 4801
QY 1214 GTGGGAAATCACCGCACACTTTGACTCCATCTGTTGTAAGTAACACCTGCGCTGTTCCG 1273
DB 4800 GCTAGCTGTGACTCCCTACTTTTGTATTTCTTTATTTGGTTAAAGTCTGTACGATGTTCT 4741
QY 1274 ACTTTGAACCTGTGTGTGCTGCTGACAGCGCGGTTGGCTGAGTTCCAGCTGTCTGTTG 1333
DB 4740 CTTTGAACAGGCAATCAGTAAATGCAAGCTGCTTAAAGAAATTTCCGATTCGTCGCG 4681
QY 1334 TGCACCAACCAATTTGTTGCTGCTGCTGCGGAGAGGACTTCACTTTTCCAAGC 1393
DB 4680 TGAACCAACAAATTTTCCGTTTATCAAAATGTTGTGAGCTATCCAGCGTTTCAATCTCGCG 4621
QY 1394 GCATCGCCACCGGATTTATCGCGCATCACCCACACCTCTTTCAGGCTCCACCTGCGGATG 1453
DB 4620 AAGCAAAACAACTTTATTTGATTAATACACCTGAATTTTGTGAATTTTCTCGTATGCGCG 4561
QY 1454 ATGACGAGGAGCGCATCTGATTTACTTGGCAGATGTACCGTGAAACAAGCTCATGCTG 1513
DB 4560 ATCGTGGCAATAAAACAGATGAATACATTTGAGAAGTAACGGTCAATGGGTTCCCTCGCA 4501
QY 1514 TGCGTCCAAAGG-----ATGTTGCAGCAACCAATCGATTAAGCTGCCCAACATCAAGGATC 1567
DB 4500 TTGAAGAAACGAGAAABAATAATTTTGAAGCCCGCGGTGCCGACAGATATTGAAGTTC 4441
QY 1568 TGCCACTGCCACGGGTTCCCGTGACCGCTGAGCAGCTTGGCCGACCGCGTTTCTC 1627
DB 4440 CAGAAAGTGAATTACAGCTAAATAATTTCTAGCGCTCAAGGAGCAACCGCGTCAATCG 4381
QY 1628 GTGATCTCCGTGAGAGGAGCACTGGAGTTACTGTATACACCTTCCGCGATGACAC 1687
DB 4380 ATTGGGTTAAGAAACAAGAAAGTGTGTTAATGACAGATACACCTTCCGTATGCGCAC 4321
QY 1688 AGTCTTTGCTTCCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGCGAGAGCGCGT 1747
DB 4320 AAAAGTTTATTAGTACTCTGCTGAGAAACAAGATTTTAAAGCAATTTGCTGTCTAACTG 4261
QY 1748 CAAAGCTGACTCTCTGAGCTTTTGTCCGTGAGGCTTGGGCGCGCGACCTTACGATG 1807
DB 4260 ATGACGCTCCCTGAGCTGTTTCTAGTGAATTTGGGGTGGCGCTTACGTTCTGATG 4201
QY 1808 CGATGCGTTCTCTTTGAGGATCCGTGGGACAGGCTCGACAGCTGCGCGAGCGGATG 1867
DB 4200 CCTATCGCTTCTTAACCGAAGACCACTGGCAACGTTTAAAGAAAAATTTGTCAGTTGATG 4141
QY 1868 CGAATGTAAACATTTCAGATGCTGCTTCCGCGCGCAACACCGTGGGATACACCCCGTACC 1927

Db	4140	CAACACACACTTTTACAAATCTCTGTTACAGAGGATCTAAATGCGGTAGGCTATCAAAATTATC	4081
Qy	1928	CAGACTCCGCTC GCGCGGCTTTGTTTAAGGAAGACTGCAGAGCTCCGCGCTGGACACTCTTCC	1987
Db	4080	CTGATAATGTTATTGAAGAANTTTATTAAAGAATCCGCTCGCCAAGAGTCGATGTTTCC	4021
Qy	1988	GCATCTTCGACGCGCTTAACGACGCTCTCCAGATGGTCCAGCAATCGACGAGTCTCG	2047
Db	4020	GAATTTTTTGATAGCTTAAACTGGAGTTCTCNAATGGAAAAAGTAGTAATCAAGTCGTTCCGG	3961
Qy	2048	AGACCAACACGCGGTAGCGAGTGGCTATGGCTTAATCTGGTGATCTCTCTGATCCAA	2107
Db	3960	ATACCGGAA--AAATGCGGAACAGCAATTTGTTATCTGGGACATCAATGATCCAG	3904
Qy	2108	ATGAAAAGCTCTACACCTGGATTACTACCTAAAGATGGCAGAGAGATCGTCAAGTCTG	2167
Db	3903	CCGAGCAAAATATAATGTTTCAATATTAATCTTGATATGGCTAAAGAAATTTGAAAAATTTAG	3844
Qy	2168	GCGCTCACATTTCTGGCCATTAAAGGATATGCGCTGGTCTGCTTCGCCCCAGCTCGGCTAACCA	2227
Db	3843	GTGCACAATCATTTGCGATTAAGATATGCGCTGGCTTATGAAACCACAAGCTGCTTATC	3784
Qy	2228	AGCTGTTACCGCACTCGCGCGTGAAATCGATCTGCCAGTGCACTGACACCCACGACA	2287
Db	3783	GTTTAAATTAGTGAATTAAGAGCGCAGCCAGATTTACCAATTCATCTCCACACTCAACGACA	3724
Qy	2288	CTGCGGTGCGCAGTTGGCTTACTTCTTCTGCGAGCTCAAGCTGGTCAGATGCTGTTG	2347
Db	3723	CTAGTGGCAATGGGATCATCACATATTCACGAGCTACTAAAGCGGCGGCTTGATATTGTTG	3664
Qy	2348	ACGCTGCTTCGCGACCACTGTCTGGCACACACTCCACGACCATCCCTGCTGCCACTGTTG	2407
Db	3663	ACGTCGCAATGAGTGCAGTAGTGGCGCAACTAGCCAGCAAGTATGAACAGTCTATATT	3604
Qy	2408	CTGCATTCGCGCACACCGTTCGGGATACCGGTTTGAGGCTTGAGGCTGTTTCTGACCTCG	2467
Db	3603	ATGCTTTTAGTCAATGGTGAAACGGACGCCAACTATTAACTTGAATATGACACAGAAAAATCA	3544
Qy	2468	AGCGCTACTGGGAAGCTGTGCGCGGACTCTACTGCGCAATTTGAGTCTGGAACCCACAGCC	2527
Db	3543	ATCATTTATTGGGAAGATGTGGCATGTAATTATCAACCATTTGAAAATGGGTTAAATGCC	3484
Qy	2528	CAACCGTTCGCGTCTACCGGCACGAAATCCCAAGCGGACAGTTGTTCACCTCGGTGCAC	2587
Db	3483	CGCAAAACAGAAGTCTATATGCATGAATCCCTGGTGGGCAATCTCTAATCTACAGCAGC	3424
Qy	2588	AGGCCACCGCACTGGGCTTGTGATCGCTTCGAGCTCATCGAGACAACTACGACGCG	2647
Db	3423	AGCAAAAGCGTAGGCTTAGGCAACCGTTGGGATGAATCAAAAAATATATCACACAG	3364
Qy	2648	TTAATGAGATGTGGGACGCCCAACCAAGGTCAACCCCATCTCTCAAGGTTGTTGGCGACC	2707
Db	3363	TGAATTTGATGTTTGGCGATATTGTTAAAGTGACGCCATCTTCAAAAGTCGTGGGAGACA	3304
Qy	2708	TCGCATCCACCTGTGTGGTGGGGTGATAGATCCAGCAGACTTCTGTCAGACCCACAA	2767
Db	3303	TGGCACTCTTCATGGTTCAAAATAATCTGACAGAACAGATGTTTATGCGCGTGGTGAAG	3244
Qy	2768	AGTACGACATCCACGACTCTGTGTCATCGCGTTCCTCGCGGGCGAGCTTGGTAACCTCCAG	2827
Db	3243	AACATAAGTTTTCTCGAATCAGTTGTGACTTTTTTCCAAAGTGATTTAGGTCAACCAAGTTG	3184
Qy	2828	GTGGTGGCCAGAACCACTGC	2848
Db	3183	GTGGGTTCCGGAAGAGCTGC	3163

RESULT 10

US-09-815-242-4413

00 00 013 242 4413
: Sequence 4413, Application us/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

```

; APPLICANT: Haseelbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 4413
; LENGTH: 3441
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4413

```

Query Match 14.0%: Score 486.8: DB 10: Length 3441:

Best Local Similarity 48.4%; Pred. No. 2.5e-140;

Matches	Conservative	Mismatches	Indels	Gaps
1664	0	1712	60	9

	Qy	84	ATTCAAAGAAGTCTTGTTAGCAAAACCGCGGGAATTCGGGTCCGGTCTTCCGTGCAGC	1439
	Db	6	AATAAAAAAAGGTACTTGTCTAACCGTGGAGAAATTGCCAATTCGTATATTCAGAGCGGC	65
	Qy	144	ACTCGAAAACGGTGCAGCCACGGTAGCTATTATACCOCGGTGAAGATCGGGGATCATTTCCA	203
	Db	66	AGCAGAATTAGACATCAGCACAGTTGCAATTTATTGCAATGAAGACAAGAAGTTCATTACA	125
	Qy	204	CCGCTCTTTTTCTCTGAAGCTGTCCGCAATTGGTACTGAAAGGCTCACAGTCAAGCGGTA	263
	Db	126	TAGATAAAAAAGCAGATGAATCCCTATTATTAGTTGGAAGTGATTTTAGGTCTCTGCTGAAAAGTTA	185
	Qy	264	CCTGGACATCGATGAANAATTATCGTGCAGCTATAAAAAGTTTAAAGCAGATGCTATTTTACCC	323
	Db	186	TTTAAATATTGAGCGTATCATGATGTAGCAAAACAAGCGAATGTGGATGCGATTCATCC	245
	Qy	324	GGGATATGGCTTCTCTGTCGAAAATGCCAGCTTGGCCGCGAGTGC CGGAAAAACGGCAT	383
	Db	246	TGGCTATGGAATTTTAAAGTGAAAATGNAACAATTTGGCGTCTGTTGTGCAAGAAGGGAT	305
	Qy	384	TACTTTTATTGGCCCCAACCCGAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGTGGGT	443
	Db	306	TAAATTTATTGGTCTCATTTTAGAACATTTAGATATGTTTGGAGATAAAGTTAAAGCTCG	365
	Qy	444	AACCGCGCAGAGAGAGCTGGTCTGCCAGTTTTGGCGGAATCCACCCGAGGCAAAAACAT	503
	Db	366	TACAACGGCTATCAAGGCGAGATTTACCAAGTATTCTCTGGTACAGACGGTCCAAATTAATC	425
	Qy	504	CGATGACATCGTTTAAAAGCGCTGAAGSCCAGAC- - -TTACCCCATCTTTGTGTAAGGCAGT	560
	Db	426	ATATGAATTTAGCAAAAGAAATTTGCAGNAGAGCTGGTTTTCCCCTAATGATTAAGGCAC	485
	Qy	561	TGCCGGTGGTGGCGGACGCGGTATGGCTTTGTTCTTCACTGATGAGCTCCGCAAAAT	620


```
Qy 2757 AGACCCACAAAGTACGACATCCAGACTCTGTCTATCGCGTTCTGCGCGGAGCTTGG 2816
|||
Db 2700 AGATGGCTATAAATTTAGATTTCCAGAAATCAAGTATGCTGCTTCTTCAAGGTTGAATAGG 2759
|||
Qy 2817 TAACCTCCAGGTGGCTGGCCAGAACCACTCGCACCGCGCACTGGGAAGCCGCTCGGA 2876
|||
Db 2760 ACAACCTGTAAATGGTTTAAATAAGATTTTACAAGCGGTATTATTTAAAGGCCAAGAGC 2819
|||
Qy 2877 AGCAAGGACCTCTGACGGAAGTTCTTGAGGAAGAGCAGCGCACCTTCGACGCTGATGA 2936
|||
Db 2820 ACTAACAGCTCGTCCAGGTGAATATCTAGAGCCGCTGATTTTGAAGAAAGTCCGTGAGTT 2879
|||
Qy 2937 TTCCAAGGAACGTGCGACACGCC-----TCACGCCCTGCTGTT 2975
|||
Db 2880 GCTTGAAGAGAGCAACAAAGTCTCTGTACGGAAGCAAGATATTATGTTATGTTATA 2939
|||
Qy 2976 CCCGAAGCCAAACCAAGAGTTCTCGAGCACCGTCGCGCTTCGGCACACCTCTGCGCT 3035
|||
Db 2940 TCCAAGATATAGACATATATTCAAACTAGAAATCAATACGGAACCTTATCGTTACT 2999
|||
Qy 3036 GGATGATCGTGAATCTTCTAGGACTGGTCGAGGGCGCGAGACTTTTGATCCGCTGCC 3095
|||
Db 3000 TGATACGCTACATCTCTTTGGAAATGCGTAATG---GTGAACAGTAGAAATCGAAAT 3056
|||
Qy 3096 AGATGCGGACCCCACTGCTTGTTCGCTCGATGCGATCTCTGAGCAGACGATGAAGG 3155
|||
Db 3057 CGATAAAGGTAAACGATTAATTTAACTAGAAACGATTTAGTGAACCCAGATGAAATGG 3116
|||
Qy 3156 TATGCGCAATGTTGGCCAAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGACCG 3215
|||
Db 3117 TAATGAAACGATTTACTATGCGATGATGTGCAAGCGAGAGCTATTTATATTAAGATGA 3176
|||
Qy 3216 CTCGTTGAGTCTGTCAACCGCAACCGCAGAAAAGGCGAGATTCTTCCAAGAGGCGCCATGT 3275
|||
Db 3177 AATGTGCATACAAATGCGAAGCTTAAGCCAAAGCAGATAAGAGTAATCCAAGTCATAT 3236
|||
Qy 3276 TGCTGCACCATCGTGGTGTGTGTCACAT---GTGACTGTTGCTGAAGTGTAGGTGCAA 3332
|||
Db 3237 CGGTGCGCAATGCGCGGTTCCAGTAACCTGAAAGTCAAGGTTAGTGTAGGTGAACTGTGAA 3296
|||
Qy 3333 GGCTGGAGATCGAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTC 3392
|||
Db 3297 AGCTAATCAGCGTGTGCTAATTTACTGAAGCTATGAAATGGAACCAACAATCAAGCACC 3356
|||
Qy 3393 TGTTGACGCGAAGATTGAACCGTTGTGTTCTGCTGCAACGAAGGTGGAAGGTGGCGA 3452
|||
Db 3357 ATTTGACGGTGTATTAACAAGTAACCTGTAAATAATGTTGATCAATAGCGACAGGCGA 3416
|||
Qy 3453 CTTGATCGTCTGTT 3468
|||
Db 3417 TTTATTAATCGAAATTT 3432
|||
```

RESULT 11

```
US-09-815-242-8263
; Sequence 8263, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haeelbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
```

```
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8263
; LENGTH: 3222
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3222)
US-09-815-242-8263

Query Match 12.3%; Score 428.8; DB 10; Length 3222;
Best Local Similarity 48.1%; Pred. No. 2.6e-122; Indels 60; Gaps 9;
Matches 1543; Conservative 0; Mismatches 1607;

Qy 310 GATGCTATTTACCGGGATATGGCTTCTCTGAAATGCCAGCTTGGCCGCGAGTGC 369
|||
Db 4 GATGCGATTCATCTGCTATGATTTTAAAGTGAATTAAGATTTGGCGTCTGT 63
|||
Qy 370 GCGGAAACCGCATTAATTTATTTGGCCCAACCCAGAGGTTCTTGATCTCACCGTAT 429
|||
Db 64 GCAGAAGAGGGATTAATTTTATTTGGTCTCTATTAGAACATTTAGATATGTTTGGAGAT 123
|||
Qy 430 AAGTCTCGTGGTAAACCGCGCGGAGAGAGCTGCTGCGAGTTTGGGGAATCCACC 489
|||
Db 124 AAAGTTAAAGCTCGTCAACCGGCTATCAAGGCGAGTTTACCAGTTATTCTCTGGTACAGAC 183
|||
Qy 490 CCAGAGCAAAAATCGATGACATCGTTA---AAAGCGCTGAAGGCCAGACTTACCCCATC 546
|||
Db 184 GTTCCAATTAATCATATGATTAAGCAAAAGATTTGCAGAGAGCTGGTTTCCGCTA 243
|||
Qy 547 TTTGTAAGCGAGTGTCCGCTGGTGGCGGACGCGGTATGCGTTTGTTCACCTGAT 606
|||
Db 244 ATGATTAAAGCCCAAGTGGTGGCGGCTAAAGGTATGAGAAATCGTTCTGTGAAGAAAT 303
|||
Qy 607 GAGTCCGCAAAATGGCAACAGAGCATCTGTTGAAGCTGAAGCGGCAATTCGGCGAGGT 666
|||
Db 304 GAATTAGAAGATGCTTTTCCATAGAGCAAAAATCAGAACAGAAAATCATTTGGTAAATAGT 363
|||
Qy 667 TCGGTATATGTCGAACGTGCTGATTAAACCCAGCACATTTGAAGTGCAGATCTTTGGC 726
|||
Db 364 GAAGTTATACATAGAAAGATACATTGATATCAAAGCATTTGAAGTACAGTACATAGT 423
|||
Qy 727 GATCGCATCGGAGAAGTTGTACACCTTTTATGAACGTGACTGCTCACTGACGCGTCGAC 786
|||
Db 424 GACGAACATGAAATATCGTACACTTATTTGAAGCTGATTGTTTCACTACAGCTCGTCAAT 483
|||
Qy 787 CAAAAGTTGTGAAATTTGGCCGAGCAGACATTTGGATCCAGAACTGCGTGTATCGCAT 846
|||
Db 484 CAAAAGTTGTGAAAGTTGCACCATCAGTTTGGATTATCACCACACATTTACGTCACAGTATT 543
|||
Qy 847 TGTGCGGATCGATTAAGTTCTGCGCTCATTTGGTTTACCAGGCGCGGAACCGTCGAA 906
|||
Db 544 TGTGATGCTCAATTCATTTGATGGAATAATTAATATGTCATGCGAGTACTGTTGAA 603
|||
Qy 907 TTCTTGTGATGAAAGGCAACCAAGTTTTCATCGAAATGAACCCACGCTATCCAGGTT 966
|||
Db 604 TTCTAGTATCTG---GTGACGAATCTCTTTATAGAAAGTTAACCTCGTGTACAGTA 660
|||
Qy 967 GAGCACACCGTGACTGAAGAAAGTCAACCGAGGTGACCTGTTGAAGGCGGAGATCGGTTG 1026
|||
```

Ds 661 GAGCATCAAAATTACAGATGGTAAAGAAATTGATATTGTTAAGACACAAAATTTAGTT 720
Qy 1027 GCTGCTGGTCAACCTTTGAAGGAATTGG-----GTCTGACCCCAAGATAAGATCAAG 1077
Ds 721 GCAGCAGGTGCGGATTTATTTGGTGAAGAAATTAATATGCGGCAACAAAAGATATTACA 780
Qy 1078 ACCACCGGTGAGCAGCTGCGAGTCCGATCACCACGGAAGATCAACACACGCTTCCGC 1137
Ds 781 ACATTTGGCTATGCCATCCAAATGTCGATTACAACAGAAATCGTTAAATGATTTATG 840
Qy 1138 CCAGATACCGGAATATCACCGGTACCGCTCACAGCGGAGTGGCGTTGCTGTTGAC 1197
Ds 841 CCGATACTGGAACAAATCAITGCTTATCGTTCAAGCGTGGCTTTGGTGAAGTCTTGAT 900
Qy 1198 GGTGAGCTCAGTCC---GGTGGGAATCACCGCACATTTGACTCCATCGTGGTGA 1254
Ds 901 GCTGGAATGTTTCCAAAGTGTGAGATATCACTTTATTAATCAITTAATCTGTA 960
Qy 1255 ATGACCTGCGGTGTTCCGATTTGAACTGCTGTTGCTCGTGACACAGCGGTTGGCT 1314
Ds 961 TTATCTACATCGATATCATTTAAACAGCAGAAATAAATGGTACGTTCAITACGA 1020
Qy 1315 GAGTTACCGTGTCTGGTTGCAACCAACATTTGTTTCTGCGTGGTGTGTCGCGGAA 1374
Ds 1021 GAAATGCGTATTCGTGGTTTAAACCTAAATTTCCATTTTAAATTAATGAAGAAT 1080
Qy 1375 GAGGACTTCACTTCCAAAGCGATCGCCACCGGATTTATCGCGGATCACCCACACCTCTT 1434
Ds 1081 AAAAGTTTCAACAGTGTGATTAACAACCTAAATTTATTTGAAGAAACACCAAGAACTTTT 1140
Qy 1435 CAGGCTCACCTGCGGATGATGACGAGGACGATCTCGGATTAATTTGCGAGATGTCA 1494
Ds 1141 GACATTCAGCGCTCTCTAGATAGAGGTACTAAACATTTAGATATATATAGTAAATGA 1200
Qy 1495 GTGAACAGCTCATGTGT-----GGTCCAAAGGATTTGCGACCAACATC 1542
Ds 1201 ATTAATGTTTCCAAATGTTGAGAAAGTCCGAAACCGACTATGAATTTAGCATCAAT 1260
Qy 1543 GATAAGTGCACACATCAAGATCTGCCACTGCCACGCGGTTCGCGTACGACGACGCTGAG 1602
Ds 1261 CCAATGATCTTCAAGTAAATCGCTTCAIT---TAGTGGTACGAAACAAATTTGCTGAT 1317
Qy 1603 CAGTTGGCCGAGCGGTTGCTGATGATCTCGTGACGAGGACGACGCTGAGTACT 1662
Ds 1318 GAAGTAGTCCAAAGGTGTAGCTGAATGGGTTAAAGCAGGATGATGTTCTACTAACA 1377
Qy 1663 GATACCACTTCCGCGATGCAACAGTCTTTGTTGCGGACCGAGTCCGCTCAITTCGA 1722
Ds 1378 GATACCACTTTAGAGTGCACCAATCAATTAATTAATGATACAGAGTTAGAACTAAGGAT 1437
Qy 1723 CTGAAGCTCGGCGAGGCGCTCGCAAGCTGACTCTGAGCTTTTGTCCGTGGAGGCC 1782
Ds 1438 ATGATTAATATCGCATCCAAACAGCGGACGTAATTAAGATGGTTTCTCACTAGAAATG 1497
Qy 1783 TGGGCGCGGACCTACGATGTGGCGATGCTTCTCTTTAGGATCCGTGGGACAGG 1842
Ds 1498 TGGGCGGTGCTACATTTGATGGCATATAATTTCTTGAAGGAAACCCATGGGAACGA 1557
Qy 1843 CTGACGAGCTGCGGAGGCGATCCGAATGAAATTAACATTCAGATGCTGTTCCGCGCGC 1902
Ds 1558 CTGGAAGCTCTACGTAAAGCTATTTCCAAATGATTAATTTCCAAATGTTGTTACGTGCTCA 1617
Qy 1903 AACACGTGGGATACACCCCTACCGAGCTCCGCTGCGCGGCTTTGTTAAGGAGCT 1962
Ds 1618 AACGAGTTGTTTAAACATCTCTGATTAATGTTTATTAATAATTCGTACAGAAATG 1677
Qy 1963 GCGAGCTCGCGTGGATCTTCCGATCTTCCAGCGGCTTAAGACGCTCCAGATG 2022
Ds 1678 GCTAAGCAGGATAGATGCTTTAGAAATTTTCGATTTCAATTAACCTGGGATGAATG 1737
Qy 2023 CGTCAGCAATCGAGCTGCGAGGCCAACACCGGGTAGCCGAGGTGGCTATGGCT 2082
Ds 1738 AAAGTTGCCAATGAAGCAGTAAAGGTAACGATTAATTTAATACTAGAACGATTAAGTGA 1797

Qy 2083 TATTCTGTGATCTCTCTGATCCAAATGAAGGCTCTACACCTGGATTAACCTAAAG 2142
Ds 1798 ACAGTGCATTTTAAATCTCTGAGCGATCAAAATTTATATCTTTAGAGTATTTATGCAAA 1857
Qy 2143 ATGCGAGAGGAGTCTGCAAGTCTGGCGCTCAATTTGCGCAATTAAGGATATCGCTGGT 2202
Ds 1858 CTAGCTAAAGAGTTAGAACGTTGAAGGTTTCCATATTTTAGCGATTAAGATATGCGCAGC 1917
Qy 2203 CTGCTTCGCCAGCTGGGTAAACCAAGCTGTCACCGCACTGCGCGCTGAATTCGATCTG 2262
Ds 1918 TTATTAACACCTAAAGCTGCTTACGAATTTGATTTGATTTAAAGCAGCTGTAGATTTA 1977
Qy 2263 CCAGTGCAGCTGCACACCCAGCACTGCGGTGCGCACTGCTACTCTTTGCTGCA 2322
Ds 1978 CCAATCCATCTTCACACTCATGATACAAGTGTAAATGTTTATTAACGTAACAAAGCA 2037
Qy 2323 GCTCAAGCTGTCGAGATGCTTTGACGGTCTTCCGACCACTGCTGGCACCACTCC 2382
Ds 2038 ATAGATGCTGGTGCATATCATTTGATCTGCTGTTCAATGATGGTTTAAACAAGT 2097
Qy 2383 CAGCATCCCTGCTGCAATTTGCTGCTGCTGCGCACACCCGTCGGGATACCGGTTG 2442
Ds 2098 CAGCCACGCGCAATTCGTTATATTAATGCTTTCCACGCTCACTTAGAACT 2157
Qy 2443 AGCTCGAGGCTGTTTGAACCTCGAGCGTACTGGAAGCTGTGCGGAGCTGTACCTG 2502
Ds 2158 GATATTTGAAGTATGGAGTCACTTAGTCAATTTATTTGGTCACTGACGCTATTTATTTCA 2217
Qy 2503 CCAATTTGAGTCTGGAACCCCGAGGCCCAACCGGTCGCTTACCGCCACGAAATCCAGGC 2562
Ds 2218 GACTTTGAAGTATGATCAAAATCACCGAATACTGAAATTTATCAACATGAATGCTGGT 2277
Qy 2563 GGACAGTTGTCCAACTGCTGTCAGCGCCACCGCACTGGGCTTGTGCTGCTGCTGAG 2632
Ds 2278 GGACAGTATTCGAATTTAAGTCAACAAGCTAAAGTTTAGGTTAGCGCAAGATTTGAT 2337
Qy 2623 CTCATCGAAGCAACTACGACGCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCA 2682
Ds 2338 GAAGTCAAGATATGATCGCAGAGTGAATTTCTTATTTGCTGATATCGTAAAGTAAAG 2397
Qy 2683 CCATCTCCAGAGTTGTTGGCGACTCGCACTCCACCTGTTGTTGGGGTGTAGATCCA 2742
Ds 2398 CCATGCTTAAAGTATTTGGTATATGCACTTTATATGTTACAAATGATCTTGATGA 2457
Qy 2743 GGAGCTTTGCTGCGAGCCCAAAAGTACGACATCCGAGACTGTGATCGGTTCCG 2802
Ds 2458 CAATCCGATTAACAGATGGCTATAATTAATTTCCGAAATCAGTAGTGTGTTCTTC 2517
Qy 2803 CGCGCGAGCTTGTGTAACCTCCAGTGGCTGGCGAGCAACCACTGCGCACCCGCGCACTG 2862
Ds 2518 AAAGTGAATAGGACAACTGTAAATGGTTTAAATAAGATTTTACAGCGGTTATTTTA 2577
Qy 2863 GAAGCCCTCCGAAGGCAAGCACTTCTGACGGAAGTTCTGAGGAAGAGCAGCGCAC 2922
Ds 2578 AAAGCCGAAGAGCACTAACAGCTCGTCCAGTGAATATCTAGAGCGGTTGATTTTGA 2637
Qy 2923 CTCGAGCTGTGATTTCAAGGAACGTCGCAACAGCC-----TC 2961
Ds 2638 AAAGTCCGTGAGTTGCTTTGAAGAGAGCAACAAGGTCCTGTTACGAGCAAGATATTTAT 2697
Qy 2962 AACCCCTGCTGTTCCCAAGCAACCGAAGTTCTTCGAGCACCGCTCGCGCTTCGCG 3021
Ds 2698 ACTTATGATTTATATCCAAAAGTATATGAACAATATTTCAAACTAGAAAATCAATACGA 2757
Qy 3022 AACACCTGCGCTGGATGCTGAATTTCTTCTAGCACTGCTGCGAGCGCGCGAGAT 3081
Ds 2758 ACTTATGTTACTTGTATACGCTACATTTCTTTTGAATGCGTAAATGTT---GAAACA 2814
Qy 3082 TTGATCCGCTGCGAGATGTCGCAACCACTGCTGTTGCGCTGGATGCTGCTGAG 3141
Ds 2815 GTAGAAATCGAAATCGATAAAGGTAACGATTAATTTAATACTAGAACGATTAAGTGA 2874

Db 959 ----AAGCGCTGGCTTCGGCAAGCCGATATTAATCACTGGCGCCATCGATCGAATGC 1014
Qy 1102 CGCATCACCGAAGATCAAAACAACCGCTTCGGCCAGATACCGGAATATCACCGCG 1161
Db 1015 CGTATCAATCGGAAGATCGGAATACAATTCGGCGGAATCCTGGCGCATTCACAGC 1074
Qy 1162 TACCGCTCACCGCGGAGCTGGGTTCTGCTTGAAGGTCAGCTCAGCTCGGTGCGAA 1221
Db 1075 TATTTACCGCGCGCGCCCGGCTTCTGCTGATTCCTATCCAGCTAGAA 1134
Qy 1222 ATCAACCGACATTTGACTTCATCTGGTGAATATGACCTGCGGTTCGACTTTGA 1281
Db 1135 ATTCCGCCCTTACGATTCGCTATTTGGCAATTTGATCTCGGGTGCAACCGGAA 1194
Qy 1282 ACTGCTTGGCTGTGCACAGCGCGCTTGGCTGAGTTACCGTGTCTGTTGCAACC 1341
Db 1195 GAGCGATCGCGGATGACAGCGTCTCGCGGAATGCGGCATACCGCGCTTGCAGC 1254
Qy 1342 AACATGCTTCTTGTGCTGCTGCGGGAAGAGACTTCTCTCCAAAGCGCATCGCC 1401
Db 1255 ACCCTTAGTTTCCATCAGCTGATGTTGACAGATGCTGAGTTCTCGCGCGGAACCTCTAT 1314
Qy 1402 ACCGATTTATCG 1414
Db 1315 ACCAACTTTGTTG 1327

RESULT 13

US-09-815-242-6233
; Sequence 6233, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Olsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 1410
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6233
; LENGTH: 1350
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1350)
US-09-815-242-6233

Query Match 8.7%; Score 302.2; DB 10; Length 1350;
Best Local Similarity 54.9%; Pred. No. 3.4e-83;
Matches 714; Conservative 0; Mismatches 563; Indels 24; Gaps 5;

Qy 89 AAAAGATCTTTGGTAGCAAAACCGCGCGAAATTCGGGTCGGTTCCTTCGTGACGACTCG 148
Db 8 ATAAAAATTTGTTATTTGCCAACCGCGCGAGATTTGATTTGCGTTATTTCTTCGTGCTGTAAAG 67
Qy 149 AACCGGTGACGCCACCGTAGCTATTTACCCCGTGAAGATCGGGGATCATTTCCACCGCT 208
Db 68 AACTGGGCATCAAGACTGTCGTGTGCACTCCAGCGCGATCGGATCTAAAAACAGTAT 127
Qy 209 CTTTGTGCTTCTGAGCTGTCGGCATTCGTTACTGAAGCTCACAGTCAAGGCGTACCTGG 268
Db 128 TACTGCGAGATGAACCGTCTGTATTTGSCCTGCTCGT---CAGTAAAGTATCTGA 184
Qy 269 ACATCGATGAATTTATCGGTGACGCTAAAAAGTTAAAGGAGATGCTATTTCACCGGAT 328
Db 185 ACATCCCGGCAATCATCAGCGCCCTGAATCACCAGCGCAGTAGCAATTCATCCGGGT 244
Qy 329 ATGCTTCTCTGTGTAATATCCCGCTTTCGCCGAGTGGCGGAAAAACGCAATTAATT 389
Db 245 ACGGCTTCTCTCCGAGAACGCCAACTTTGCGAGCAGGTTGAACGCTCCGGGCTTTATCT 304
Qy 389 TTATTGSCCAACCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCTGCGGTAAACCG 448
Db 305 TCATTGGCCCGAAAGCAGAAACCATTCGCTGATGGCGCAAAAGTATCCGCAATCGCG 364
Qy 449 CCGCAAGAGAGGCTGTGCTGCGAGTTTGGCGGAATC-----CACCCCGAGCAAAAAACA 502
Db 365 CGATGAAAAAAGCGGCGTCCCTTTCGCTACCGGTTCTGACGGCCCGCTGGCGCAGATA 424
Qy 503 TCGATGACATCGTTAAAAAGCGCTGAAGCGCCAGACTTACCCCATCTTTGTAAAGCAGTTG 562
Db 425 TGGATAAAAACCGTGCCATTGCTAAACGCATTGTTATCCGGTGATTTCAAAAGCCCTCG 484
Qy 563 CCGGTGTGGCGGACGCGGTATGCGCTTTGTTTCTTCCACTGATGAGCTCCGCAAAATGG 622
Db 485 CCGCGCGCGCGGCTGCGGTATGCGGTAGTGGCGGCGACGCTGAACGGCACATCCA 544
Qy 623 CAACGAAGCATCTCTGTAAGCTGAAGCGCATTCGCGGACGTTCCGTATATGCAAC 682
Db 545 TCTCATGACCCGTCGGAAGCAAGTCTCTTCAGCAAGATATGTTTACATGAGA 604
Qy 683 GTGCTGATTAACCCCGCAGCACATTAAGATGACATCTTTGGCGATCGCACTGGAGAAG 742
Db 605 AATACCTGGAATAATCCTCGCCACGTCGAGATTCAGGTACTGGCTGACGGTCAGGGCAACG 664
Qy 743 TTGTACACCTTTTATGAACGTGACTGCTCACTGACGCTGCTCAACAAAAAGTTGTGAAA 802
Db 665 CTATCTATCTGGCGAAGCGTGAATGCTCCATGCAACCGCCACACAGAAAGTGTGCGAAG 724
Qy 803 TTGGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATCGAGTAA 862
Db 725 AAGCGCCAGCACCGGCAATTACCCCGGAATCGCTGCTATATCGGCAACGTTGCGCTA 784
Qy 863 AGTTCTGCGCTCCATTGTTTACAGGGCGCGGAAACCGTGGAATTTCTTGGTCCATGAAA 922
Db 785 AAGCGTGTGTTGATATCGGCTATCGCGTGCAGGTACTTTTCGAGTTCTCTGTTTC---GAAA 841
Qy 923 AGGCGAACACGTTTTCATCGAATGAAACCGCTATCCAGTTTGAGCACACCGTGAAGT 982
Db 842 ACGGCGAGTTCTATTTTCAATGAAATGAACCCGCTATTCAGTTAGAACACCCCGTTACAG 901
Qy 983 AAGAAGTCAACGAGGTGGACCTGTGAAGCGCAGATCGCTTGGCTGCTGCTGCTGCACT 1042
Db 902 AATGATCAACCGCGTTGACCTGATCAAAAGAACAGTCTGCTATCGCTGCGGTCACCGC 961
Qy 1043 TGAAGGAATTTGGTCTGACCCCAAGATGAATCAAGACCCAGGTCGAGCACTGCAAGTCC 1102
Db 962 TGTGCG-----ATCAAGCAGAAAGAAAGTTTACGTTTCGCGGCCATTCGCGTGAAGTGC 1012
Qy 1103 GCATCACCAACGAAAGATCAAAACCAACCGCTTTCGCCCGCAGATACCGGAATCTATCAAC 1162
Db 1013 GTATCAACCGCGGAAGATCCGAACA---CCTTCTGCGCAAGTCCGGGCAAAATCATCCCGTT 1069

QY 1163 ACCGTCACAGCGGAGCTGGCGTTGCTTGTGACGGTGCAGCTCAGCTCGGTGGCGAAA 1222
Db 1070 TCCACGACCTGGCGGTTTGGCGTACGTTGGGAGTCTCATATACGCGGGCTACACCG 1129
QY 1223 TCACCGCACCTTTGACTCCATGCTGGTGAATAAGACCTGCGGTGTTCCGACTTTGAAA 1282
Db 1130 TACCGCGTACTATGACTCAATGATCGTAGCTGATTTGCTACGGTGAAACCGTGAAG 1189
QY 1283 CTGCTGTTGCTGTCGACAGCGCGGTGCTGAGTTCACCGTGTCTGTTGTTGCAACCA 1342
Db 1190 TGGCGATTGCCCGCATGAAGATCGCTGCAGGAGCTGATCATGACGGTATCAAAACCA 1249
QY 1343 ACATTGGTTTCTTCGTCGCTGCTGCGGGAAGAGGACTTC 1383
Db 1250 ACGTGTGATCTGCAGATCCGCATCATGAATGACGAGAACTTC 1290

RESULT 14

US-09-767-479-5

; Sequence 5, Application US/09767479

; Patent No. US20010036654A1

; GENERAL INFORMATION:

; APPLICANT: Haselkorn, Robert

; Gornicki, Piotr

; TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND

; METHODS OF USE

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Arnold, White & Durkee

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: Texas

; COUNTRY: United States of America

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/767,479

; FILING DATE: 22-Jan-2001

; CLASSIFICATION: Unknown

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/468,793

; FILING DATE: <Unknown>

; APPLICATION NUMBER: US SN 07/956,700

; FILING DATE: 02-OCT-1992

; APPLICATION NUMBER: PCT/US93/09340

; FILING DATE: 30-SEP-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Kitchell, Barbara S.

; REGISTRATION NUMBER: 33,928

; REFERENCE/DOCKET NUMBER: ARCD:152/KIT

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (512) 418-3000

; TELEFAX: (713) 789-2679

; TELEX: 79-0924

; INFORMATION FOR SEQ ID NO: 5:

; LENGTH: 3065 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: other nucleic acid

; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-767-479-5

Query Match 8.5%; Score 297; DB 10; Length 3065;

Best Local Similarity 54.4%; Pred. No. 2.2e-81;

Matches 691; Conservative 0; Mismatches 565; Indels 15; Gaps 4;

QY 85 TTCAAAAGATCTTTGGTAGCAACCGCGGAAAATCGCGGTCCGTCTTTCCGTGCAGCA 144

Db 1314 TTGACAAAATATAATTGCAATTCGGGAGAAATAGCGCTGCGCATTCCTCGCGCTGT 1373
QY 145 CTGAAACCGGTGAGCCAGCGTAGCTATTATACCCCGTGAAGATCGGGATCATCTCCAC 204
Db 1374 GAGGAATGGGATTGGACGATCGAGTTCATTTCGACTGTTCAGCCGGAATGCTCTTCAT 1433
QY 205 CGCTCTTTGCTGCTGAAGCTGCGCATTTGGTACTGAAAGCTCACAGTCAAGCGGTAC 264
Db 1434 GTCCAACTTCTGACGAAGCGGTTTGTATTGGC---GAACTGCTAGCGCTAAAGTTAT 1490
QY 265 CTGACATCGATGAAATATTCGGTGCAGCTAAATAAGTTAAAGCAGATGCTATTATCCCG 324
Db 1491 TTGAATATTCCTCAATATATTGCTGCGGCTTTAACGCGCAATGCGTGTATTATCTCT 1550
QY 325 GGATATGGCTTCTGCTGAAATATGCCAGCTTCCCGCGAGTCGCGGAGAAACCGCATTT 384
Db 1551 GGGTATGGCTTTTATCTGAAATATGCCAAATTTGGGAAATCTGCTGACCATCACATT 1610
QY 385 ACTTTATTTGGCCCAACCCAGAGGTTCTTGATCTACCGGTGATAAGTCTCGTGGCGTA 444
Db 1611 GCATTTCATTGGCCCAACCCAGAGGTTCTCCAGTTTGGCGG--AATCCACCCGAGCAAAACA 502
QY 445 ACCGCGCGAAGAGGCTGCTGCCAGTTTGGCGG--AATCCACCCGAGCAAAACA 502
Db 1671 GAAACCATGCAAAAGCTGCTGACCGACAGTACCGGGTAGTGAAGTTTGGTAGAGACA 1730
QY 503 TCGATGACATCGTTAAAGCGCTGAAGCCA-GACTTTACCCCATCTTTGTAAGGCGAGTT 561
Db 1731 GAGCAAGAGGATTAGNACTGGCGAAGATATTGGCTACCCAGTATGATCAAGCCACG 1790
QY 562 GCGGTGTTGGGAGCGGATATGCGCTTTTGTTCCTCATGATGAGTCCGCAAAATG 621
Db 1791 GCTGGTGGTGGCGCGGGGTATGCGACTGCTGCGATCGCCAGATGAATTTGTCAAACTG 1850
QY 622 GCACAGACATCTCGTGAAGCTGAAGCGGATTCGCGACGGTTCGGTATATGTCGAA 681
Db 1851 TTCTTAGCCGCCAAGGTGAAGCTGGTGCAGCTTTGGTAAATGCTGCGCTTTATATAGAA 1910
QY 682 CGTCTGTGATTAAACCCAGCACATTTGAAGTGCAGATCTTTGGCGATCGCACCTGGAGAA 741
Db 1911 AATTTATTGAACGTCCGCCACATTAATTTCAATTTTGGCTGATATATACCGCAAT 1970
QY 742 GTTGTACACCTTTATGAACGTGCTCTCATCTGAGCGTCTGCACCAAAAAGTTGTCGAA 801
Db 1971 GTGATTCACTTGGGTGAGAGGGATTGCTCAATTCAGCTGCTAACCAAAAAGTTACTAGAA 2030
QY 802 ATTGGCGCAGCACAGCATTTGGATCCAGAACTGGGTGATCGCATTTGTGGGATGCAAGTA 861
Db 2031 GAAGCCCCAGCCCGCTTTGGACTCAGACCTAAGGGAAAAAATGGGACAAGCGCGGTG 2090
QY 862 AAGTTCTGCGCTCCATTGGTTTACAGGGCGGGAACCGTGAATTCCTTTGGTCGATGAA 921
Db 2091 AAAGCGGCTCAGTTTATCAATTTAGCGGGGCGAGTACTATCGAGTTTTCGTAGATAGA 2150
QY 922 AAGGGCAACCAACGTTTTTCATCGAAATGAACCCAGTATCCAGGTTTGAGCACACCGTGACT 981
Db 2151 TCCGGTCAGTTTACTTTATGAGATGAACACCCCGGATTCAAGTTAGAACATCCCGTAACT 2210
QY 982 GAAGAAGTACCGAGGTGAGACTGGTGAAGCGCGAGATGCGCTTGGTGTGTTGGTGAACC 1041
Db 2211 GAGATGGTTACTGGAGTGGATTATTGGTTTGAAGCAATCAGAAATTTGCCAAGGGGAAA-- 2268
QY 1042 TTGAAGGAATTTGGTCTGACCCCAAGATAAGATCAAGACCCACCGTGCAGCACTGCAGTGC 1101
Db 2269 -----GACTTAGACTAACTCAAGACCAAGTAGTTTACGCGGTATGCGATGCAATGT 2321
QY 1102 CGCATCACACGGAAGATCCAAACAAACGGCTTCGCGCCAGATACCCGAAATATCACCGCG 1161
Db 2322 CGATCAATCCGAGACCCAGACGATTTTCGCGCCAGCACCCGACCGCATTAGCGGT 2381
QY 1162 TACCGCTCACAGCGGAGCTGCGTTCGCTTTGACGGTGCAGCTCAGTCTCGGTGGCGAA 1221

Db 2382 TATCTTCCCTGGCGCCCTGGGTGCGGATTGACTCCCAAGTTTACAGGATTACCAA 2441
Qy 1222 ATCAGCGACATCTGACTCCATCTGGTGAATAATGACCTGCGGTTCGACATTGAA 1281
Db 2442 ATTCCGCCCTACTAGATTCTTAATTGGTAATTGATGGTTGGGCGCTGATCGCGCT 2501
Qy 1282 ACTGCTGCTGCTGACAGCGCGGTGGCTGAGTTACCGTGTCTGGTGTGCAACC 1341
Db 2502 ACTGCTATTAAACCGCATGAACAGCGCCCTCAGGGAATGCGCCATCACTGGATTACCTACA 2561
Qy 1342 AACATTGGTTT 1352
Db 2562 ACCATTGGGTT 2572

RESULT 15
US-09-974-300-2
; Sequence 2, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085-500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2238
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-2

Query Match 8.2%; Score 286; DB 10; Length 2238;
Best Local Similarity 48.8%; Pred. No. 4.9e-78;
Matches 939; Conservative 0; Mismatches 955; Indels 30; Gaps 5;

Qy 1564 GATCTGCCACTGCCACGCGGTTCCTGACCGCTGAAGCAGCTTGGCCAGCGCGTTT 1623
Db 309 GATCAGCAATCGAAGCGGNAACAAACAGATTCTTGATGAACGGGAGCGCTC 368
Qy 1624 GCTCGTATCTCGTGAGCAGGACGACTGGCAGTTACTGATACCACTTCGCGATGCA 1683
Db 369 GTCAATGGGTGAAGATCAAGAGAAAGTCTCTACTGATGACGTTCCCGGACGCC 428
Qy 1684 CACGAGCTTTGCTGGACCCGAGTCCGCTCATTCGCACTGAAGCTGGCGAGAGGCC 1743
Db 429 CATCAGTCATTTGCTGGCAAGAGTCAAGACGATGACCTGAAATAATCGCAATCCG 488
Qy 1744 GTCGCAAGCTGACTCTCTGAGCTTTTCTCGGTGAGGCTTGGGCGCGGACCTACGAT 1803
Db 489 ACGGCTGCGCTCTGGCCAGAGCTTTTCAGTCTTGAATGTGGGCGCGGCGACATTTGAT 548
Qy 1804 GTGGCGATGCGTTTCTCTTTGAGGATCCGTGGGACAGGCTCGACAGCTGCGCGAGGCG 1863
Db 549 GTCGCTTACCGTTTCTTGAAGAAGATCCGTGGAAAAAGGCTCGAGGAGCTCGGAAAGAA 608
Qy 1864 ATGCGGAATGAACATTCAGATGCTCTTCTCGGCGCGCAACCGTGGGATACACCCCG 1923
Db 609 ATTCCGAATACGATTTCAATGCTTTTGAGATCTTTCGATGCTCGGATACAAAC 668
Qy 1924 TACCAGACTCCGCTCGCGCGCTTTGTAAGAGCTCCAGCTCCGCGGTGGACATC 1983
Db 669 TACCTCAACATCTGATTAAGAAAGTTCTGACGAGATCGGCTCGCGCGGAAATCGATGA 728
Qy 1984 TTCGCTATTCGAGCGGCTTAACGAGCTCTCCAGATGGTTCAGCAATCGAGGATC 2043

Db 729 TTCCGCATTTTTGACAGCTTGAACCTGGGTTAAAGGGATGACGCTCGCCATTGATGACAGT 788
Qy 2044 CTGAGAGACCAACACCGCGGTAGCGAGTGGCTATGCTTATTCTGGTGTATCTCTCTGAT 2103
Db 789 CGCGAGTCAGGCA--AGCTTGCCGAAGCGGATCTGCTATACAGGGACATTTCTCGAT 845
Qy 2104 CCAATGAAGAAGCTCTACACCGCTGATTACTACCTAAAGATGGCAGAGGAGATCGTCAAG 2163
Db 846 CCGAACAGAGCAAGATTAATCTTTGAATATTTATACGTCAATGGCAAGAGCTTTGAAGCT 905
Qy 2164 TCTGGCGCTCATCTCTGGCCATTAAAGGATATGCTGCTGCTTCGCCCGAGTGGCGGTA 2223
Db 906 GCGGGGCGCATATTTCTCGGCATTAAAGATATGGCCGCGCTGTGAGACCTCAAGCGGCA 965
Qy 2224 ACCAAGCTGGTCAACGCACTGCGCGCTGAAATTCGATCTGCCAGTGCACGACACCCAC 2283
Db 966 TACGAACTGGTGTGCGGCTTTGAAAGAGACGATCGACATTCGATCCACTTTGCATACAC 1025
Qy 2284 GACACTGCGGTGGCGAGTTGGCTTACCTACTTTGCTGCTGAGCTCAAGCTGGTGCAGATGCT 2343
Db 1026 GACAGAGCGGCAACGGGTATTTTATGTATGCGAAGCGGATAGAAAGCAGCGCTCGACATC 1085
Qy 2344 GTTGACGCTGCTTCGCGACCACTGTCTGGCACCACTCCAGCCATCCCTGTCTGCCATT 2403
Db 1086 GTCGAAGTAGCGGTGAGCTCGATGCGGCTCTGACATCACAGCCAAAGCGAAGCTCGCTT 1145
Qy 2404 GTTCTGCAATTCGCGCACACCGCTGCGATACCGGTTTGGCTTCGAGGCTGTTTCTGAC 2463
Db 1146 TACATGCGCTTGAAGGATTAACCGCTCGCAGTTCATGTCTGATCGGAAGATAA 1205
Qy 2464 CTCGAGCGCTACTGGGAAGCTGTGCGGAGCTGTACTGCTGCTGCTGCTGCTGCTGCTGCT 2523
Db 1206 TAGTCTCAAAAAAAGAGTCTGTGAGAAATTAATACAGAGAGTTGAGAGCGGATGAT 1265
Qy 2524 GGCCCAACCGGTGCGCTTACCGCCACGAAATCCAGCGGACAGTTGTTCACCTGCGGT 2583
Db 1266 GCGCTCATACCGAAATTTACAGCATGAATGCTCGCGCCAAATACAGCAACCTTCAG 1325
Qy 2584 GCAGAGCCACCGCACTGGGCTTGTGCTGATCGCTTCGAGCTCATCGAAGACCACTACGCA 2643
Db 1326 CAGCAGGCTAAGGAGTGGGCTCGCGACCGCTGGAATGAAGTGAAGAAATGTACAGC 1385
Qy 2644 GCCGTTAATGAGTGTGGGACGCCAACAAAGTCAACCCATCTCAAGTTGTTGGC 2703
Db 1386 CGGCTCAACCACTGTTCGGGACATCGTAAAGGTTACGCGCTCATCTAAGTCTCGGG 1445
Qy 2704 GACCTCGCACTCCACCTGTTGTTGGTGTAGATCCAGCAGACTTTGTGCGACGCCA 2763
Db 1446 GATATGCGCTTTACATGCTGCAAAATAACCTGACGGAAGATGATATTACGAAGGGA 1505
Qy 2764 CAAAAGTAGACATCCAGACTCTGTCTATCGCTTCCTGCGCGGAGCTTGTAAAGCT 2823
Db 1506 GAATCTCTGATTTCTGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1565
Qy 2824 CAGGTGCTGCGCAGAACCACTGCGCACCGCGACTTGAAGCGCGCTCGAAGGCAAG 2883
Db 1566 CACGCGGATTTCCCTGAGAAATTCGAAAGCTGATTTTAAAGGGCAGGAGCCGATTA 1625
Qy 2884 GCACCTCTGAGGGAAGTTCTTGAGGAAGAGCAGCGCACCTCGACGCTGATGATTCGAAG 2943
Db 1626 GTACGCTGAGAACTGCTTGAACCGGTTTCATTTGACGCTATATAAAGCGGAATTTCTG 1685
Qy 2944 GAACTC-----GCAACAGCTCAACCGCTGCTGCTTC-----CCGAAG 2982
Db 1686 GAAAAGCATGTTATGAGCTTTCCGACCAAGATGCTGTGTGATATGCTGCTTTATCCGAAA 1745
Qy 2983 CCAACCGAAGAGTTCTCGACACCGCTTCGCAACACCTTCGCGCTGCTGCTGCTGCTGCTGCT 3042
Db 1746 GTGTTTACGAGTATGTCAGACAGCGGAACTGTACGCGGATATCTCTGTCTTGATACA 1805
Qy 3043 GGTGAATTTCTTACGAGCTGCTGAGGCGCGGAGCTTTGATCCGCTCCGAGATGCTG 3102
Db 1806 CCAACCTTCTCTACGCAATGACCTCGCGGAAGAAATCGAGGT---CGAGATTGAAGA 1862

QY 3103 CGACCCCACTGCTTGTTCGGCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGGC 3162
| | | | |
Db 1863 GGGAAAACGCTGATCGTCAAACTTGTCTATCGGGGAACCGCGTCCGGATGCCACGAGA 1922
| | | | |
QY 3163 AATGTTGTGGCCAAAGTCAACGGCCAGATCGCCCAATGCGTGTGCGTGACCGCTCCGTT 3222
| | | | |
Db 1923 GTCGTTTATTTTGAAGTCAAGCGCCAGCTCGCGAAGTGTGTCATTAAGACGAAGCATT 1982
| | | | |
QY 3223 GAGTCTGTCAACCGCAACCGCAGAAAAGGCAGATTCTCCAAAGGGCCATGTTGTGCA 3282
| | | | |
Db 1983 AATCGTCCGTTTCATCAAAAGGTGAAGCCGACCGTTCAAACCGCAATCATATCGCGCA 2042
| | | | |
QY 3283 CCATTCGCTGG---TGTTGCTACTGTGACTGTTGCTGAAAGTGATGAGGTCAAGGCTGGA 3339
| | | | |
Db 2043 TCTATGCCGGGAACGGTAATCAAGCTTCTTTGTAAGCAAAAGGGGACCAAGTGAAGAGGGC 2102
| | | | |
QY 3340 GATGCACTCGCAATCATCGAGGCTATGAAGATGAAGCAACAATCACTGCTTCTGTTGAC 3399
| | | | |
Db 2103 GATCATTTGATGATCAATGAAGCGATGAATGGAACGACCGTTCAAGGCTCCGTTTCA 2162
| | | | |
QY 3400 GGCAAGATTGAACCGTGTGTGTTCTCTGCTGCAACGAAGGTGGAAGGTGGCGACTTCATC 3459
| | | | |
Db 2163 GGAACCTGTTGAAACATTACGTTTACAAACGGAGAAGCCATTCAACCGGACCTTCTC 2222
| | | | |
QY 3460 GTCG 3463
| |
Db 2223 ATTG 2226

Search completed: April 5, 2003, 12:50:28
Job time : 366 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2003, 04:31:45 ; Search time 5238 Seconds
(without alignments)
10741.337 Million cell updates/sec

Title: US-09-974-973-1
Perfect score: 3474
Sequence: 1 gtagctgctacccttg.....tgatcgctgcttctctaa 3474

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

```

1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_man.*
24: em_gss_mus.*
25: em_gss_other.*
26: em_gss_pro.*
27: em_gss_rod.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	237.4	6.8	888	17	CNS074FW
2	234.4	6.7	903	14	BQ881271
c 3	221.8	6.4	859	17	AQ875053
4	220.2	6.3	637	12	BF251052
5	218	6.3	675	9	AI109467
6	213.2	6.1	744	13	BI103175

7	207.8	6.0	640	13	BI618593
8	205	5.9	715	12	BG910468
9	201.2	5.8	641	13	BJ063722
10	197.4	5.7	887	12	BF781506
11	196.6	5.7	515	9	AU202878
12	196.2	5.6	949	14	BQ752438
13	195	5.6	666	13	BI605019
14	193.2	5.6	1030	17	CNS06M6L
15	192.4	5.5	625	9	AI293242
16	192.6	5.5	508	13	BI621845
17	189.2	5.5	921	14	BQ959497
18	189.2	5.4	895	12	BF535766
19	189	5.4	611	10	BE312170
20	186.4	5.4	860	13	BI872878
21	186	5.4	915	14	BQ730755
22	185.8	5.3	599	12	BF506511
23	184.8	5.3	553	17	AQ501087
24	184.8	5.3	2403	11	AK007782
25	182.4	5.3	725	12	BG599744
26	182.2	5.2	813	12	BG761035
27	182	5.2	556	9	AI134619
28	181.8	5.2	575	9	AI389557
29	181.8	5.2	923	14	BQ734080
30	181.2	5.2	610	13	BI082302
31	181	5.2	643	12	BF499481
c 32	179.6	5.2	780	17	BH770815
33	178.6	5.1	854	13	BI146211
34	178.4	5.1	684	13	BI718188
35	177.6	5.1	632	13	BJ124629
36	176.8	5.1	609	9	AI526573
37	175.6	5.1	616	12	BG710300
38	174.6	5.0	848	17	AQ875041
c 39	173.4	5.0	856	17	CNS0749M
40	173.4	5.0	635	13	BM156619
41	173.4	5.0	794	13	BI683931
42	173.2	5.0	725	13	BI327807
43	173	5.0	652	12	BG855209
44	169.6	4.9	598	13	BJ125540
45	168.8	4.9	683	13	BM603525

ALIGNMENTS

RESULT 1
CNS074FW/c
LOCUS
DEFINITION
clone BAOB027F04 of library BAOAB from strain CLIB 210 of Kluyveromyces lactis, genomic survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE

888 bp DNA linear GSS 07-JUL-2001

GI:12211948

1 (bases 1 to 888)

Souciot J.L., Aigle, M., Artiguenave, F., Blandin, G., Bolotin-Fukuhara, M., Bon, E., Brattier, P., Casaregola, S., de-Montigny, J., Dujon, B., Durrens, P., Lepingie, A., Morente, B., Malpertuy, A., Neuveglise, C., Ozier-Kalogeropoulos, O., Potier, S., Saurin, W., Tekala, F., Toffano-Nioche, C., Wesolowski-Louvel, M., Wincker, P. and Weissbach, J.

Kluyveromyces lactis.
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Kluyveromyces.

Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies
FEBS Lett. 487 (1), 3-12 (2000)

2 (bases 1 to 888)

Bolotin-Fukuhara, M., Toffano-Nioche, C., Artiguenave, F., Duchateau-Nguyen, G., Lemaire, M., Marmeisse, R., Montrocher, R., Robert, C., Termier, M., Wincker, P. and Wesolowski-Louvel, M.

Genomic exploration of the hemiascomycetous yeasts: 11.

Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

```
BASE COUNT      172 a      281 c      280 g      170 t
ORIGIN
Query Match      6.7%; Score 234.4; DB 14; Length 903;
Best Local Similarity 57.3%; Pred. No. 3e-51;
Matches 465; Conservative 0; Mismatches 341; Indels 6; Gaps 2;
QY 1596 CTTGAACGAGCTTGGCCAGCGGTTTGTCTCGTGATCTCGTCCAGCAGGACGACCTGGC 1655
DB 18 CTTGCTCGAGAGGGGCTGAGGCTTTGTCTCGAGCTGTGCGGAACCAACCGGGGGCTGCT 77
QY 1656 AGTTACTGATACCACTTCCGCGATGACACCACTCTTTGCTTCCGACCCGAGTCCGCTC 1715
DB 78 GCTGATGGACAGACCTTCAGGAGCGCCACCACTGCTGCGCCACTGCTGTGGCGCAC 137
QY 1716 ATTGCGACTGAGCCTCGGCGAGGCGGTGCGAAGCTGACTCTGAGCTTTTGTCCGT 1775
DB 138 CCACGATCTCAAAAGATCGCCCTATGTGCGCCACAACCTTCAGCAAGCTCTTTCAGCAT 197
QY 1776 GGAGCGCTGGGGCGGCGCACTACGATGTGGCGATCGTTTCTCTTTGAGGATCGGTG 1835
DB 198 GGAGAACTGGGAGAGCCACGTTTGACGTGCGCATCGCTTCTGTATGATGCCCTG 257
QY 1836 GGACAGCTCGACGAGCTCGCGAGGGGATGCCGAATGTAAACATTCAGATGCTGCTCG 1895
DB 258 GCGCGGCTCGAGAGCTCCGGAGCTCATCCCAACATCCCTTTCAGATGCTGCTGCG 317
QY 1896 CGGCGCAACACCGTGGGATACACCCGATACCGAGACTCGTCTGCGCGGTTTGTAA 1955
DB 318 GGGGGCCAATGCTGTGGGTACACCAACTACCCAGACAACGTGTCTTCAAGTTCTGTGA 377
QY 1956 GGAAGCTGCCAGCTCCGCGCTGGACATCTTCGCGATCTTCGACGCGCTTAACGACGCTC 2015
DB 378 AGTGCCAAAGAGATGGCATGATGTCTTCGCTGTGTGATCTCCCTCAACTACTTGCC 437
QY 2016 CCAGATCGCTCCAGCAATCGACGAGCTCTGGAGACCAACACCGCGTAGCCGAGGTGC 2075
DB 438 CAACATCTGCTGGGCATGGAGCGG--CAGGAAGTCCCGAGGCGTGTGGAGGCTGC 494
QY 2076 TATGGCTATTCTGGTGATCTCTGTATCCAAATGAAAGCTCTACACCTCGATTTACTA 2135
DB 495 CATCTCATACACGGGACGCTGGCGACCCCGACCGCACCAAGTACTCACTGCAGTACTA 554
QY 2136 CTTAAAGATGCGACGAGATCGTCAAGTCTGGCGCTCACATTTCTGGCCATTAAAGGATAT 2195
DB 555 CATGGGCTTGCCGGAAGAGCTGTGCGAGCTGGCACCACCAATCTCTGTCATCAAGGACAT 614
QY 2196 GGCTGTGCTGTTCCGCCAGCTGCGGTAAACCAAGCTGCTCACCGCACTGGG---CCGTGA 2252
DB 615 GCGCGGCTGCTGAAGCCACGGCTGCACCATGCTGTGTGCTCCCTCCGGGACCGCTT 674
QY 2253 ATTGATCTGCGAGTGACGTGCAACCCAGCAACATGCGGGGTGGCGATTGGCTACTTA 2312
DB 675 CCCGCACTCCCACTGCACATCCACCCAGCACACGTCAAGGGGCGAGCGTGGCGAGCAT 734
QY 2313 CTTTGTGCGAGCTCAAGCTGGTGAGATGCTGTTGACGCTGCTTCCGCAACCACTGCTGG 2372
DB 735 GCTGGCTGTGCCAGCTGGAGCTGATGTGGTGGATGTGTCAGCTGATTCATGCTCTGG 794
QY 2373 CACCACCTCCAGCCATCCCTGCTGCGCAATG 2404
DB 795 GATGACTTCAAGCCAGCATGGGGGGCCCTG 826
```

RESULT 3
AQ875053/c 859 bp DNA linear GSS 08-NOV-1999
LOCUS V120P6 mTn-3xHA/lacZ Insertion Library, strain Y2278 Saccharomyces
DEFINITION cerevisiae genomic 5', DNA sequence.
ACCESSION AQ875053

VERSION AQ875053.1 GI:6287297
KEYWORDS GSS.
SOURCE baker's yeast.
ORGANISM Saccharomyces cerevisiae
Saccharomycetes; Saccharomycetaceae; Saccharomycetes;

REFERENCE 1 (bases 1 to 859)
AUTHORS Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,

desStages, S.A., Cheung, K.-H., Sheehan, A., Symoniatidis, D., Jansen, R.,
Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
Hager, K., Miller, P., Roeder, G.S. and Snyder, M.

TITLE Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
Gene Disruption

JOURNAL Unpublished (1999)
COMMENT Contact: Kumar A

Michael Snyder, Dept. of Mol. Cell. and Dev. Biology

Yale University

P.O. Box 208103, New Haven, CT 06520-8103, USA

Tel: 203 432 9949

Fax: 203 432 6161

Email: anuj.kumar@yale.edu

te of mTn-3xHA/lacZ insertion.

Seq primer: GGCTTCTTTCTTTGGAAGTAC

Class: transposon-tagged.

Location/Qualifiers

1..859

source

/organism="Saccharomyces cerevisiae"

/strain="Y2278 - S288C background, cir(0) rho(0)"

/db_xref="taxon:4932"

/clone_lib="mTn-3xHA/lacZ Insertion Library, strain Y2278"

/lab_host="E. coli"

/note="Vector: pHS6-Sal; A yeast genomic DNA library

without 2 micron or mitochondrial DNA was prepared in

pHS6-Sal; genomic DNA was size-fractionated (DNA of

roughly 2-3 kb in length) prior to cloning. This library

was subsequently mutagenized with a mTn-3xHA/lacZ

mini-transposon containing lacZ, URA3, and tet resistance.

"

BASE COUNT 213 a 196 c 190 g 256 t 4 others

ORIGIN

Query Match 6.4%; Score 221.8; DB 17; Length 859;

Best Local Similarity 59.8%; Pred. No. 6.9e-48;

Matches 406; Conservative 0; Mismatches 269; Indels 4; Gaps 2;

QY 680 AACGTGCTGTGATTAAACCCAGCAGACATTCGAAGTGCAGATCCTTGGCGATCGCACTGGAG 739

DB 801 AAATTTCTGGACAAGCAAAAGCATATTGAAGTTCAATTTGTGGCCGATACACCAAAA 742

QY 740 AAGTTGTACACCTTTATGAACGTGACTGCTCACTGACGCGTCTGTCACCAAAAAGTTGTG 799

DB 741 CCGTTGTTTCATCTTTTCGAAAGAGACTGTTTCGGTGCAGAGAAA-CACCAAAAAGTTGTCA 683

QY 800 AATTTGGCCAGCAGACATTTGGATCCAGAACTCGATCGCATTTGTGCGGATCGAG 859

DB 682 AAGTGGCNCAGCAAAAGACTTTACNCGTGAAGTTCGGTGAAGTTCGCGCATTTTGACATG 623

QY 860 TAAAGTTCTCGCGCTCCATTGGTTTACCAGGGCGCGGAAACCGTGAATTTCTTGTGCTGAT 919

DB 622 TTTAATTTGGCCAAAGAGTGTGGCTACAGAAATCGGGTACTCTGTAATTTCTTGTGATA 563

QY 920 AAAAGGGCAACACCGTTTTCATCGAAATGAACCCACGATCCAGGTTTGAGCACACCGTGA 979

DB 562 ACCAAATAGACACTATTTTCATTGAAATTAATCCAAAGATCCAAAGTGAACATACCATCA 503

QY 980 CTGAAGAGTCAACCGAGGTGGACCTGGTGAAGGGCGAGATGCGCTTGGCTGCTGGTCAA 1039

DB 502 CAGAAGAAATTTACCGGTATAGATATTGTGGCGGCTCAGATCCAAATTTGCGGCGCTGCT 443

QY 1040 CCTTGAAGGAATTTGGGTCTGACCCAGATGAATCAAGACCCACGCTGCAGCTGAGT 1099

DB 442 CTCTACCCAGCTGGGCGCTATTCCAGGACAAAATTTACGACTCTGTGGCTTTTGCCATT 383

```

Qy 1100 GCGCATCAGCAGGAGATCCAAACACCGCTTCGCCAGATACCGGACTATCACCG 1159
Db 382 GCGTATTACACGGAAGACCTCGCTAGAACTTCACACCATATCCGGTAGATAGAAG 323
Qy 1160 CGTACCCCTCACACGCGGAGCTGGCTTCGTCTTGACGGTG---CAGCTCAGCTCGGTG 1216
Db 322 TGTAACCGTTCTGCAAGTGGTAAATGTTAGACTGGATGGTGAACGCTATGCAAGAA 263
Qy 1217 GCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAATGACTCCGCTGGTTCGACT 1276
Db 262 CAATAATCTCACCTCATTAAGACTCAATGCTGGTCAAACTCTCATGCTCCGGTTCCACT 203
Qy 1277 TTGAAATCGTGTGCTCGTGCACAGCGCGGTGGCTGAGTTCAACGCTGTCTGGTGTG 1336
Db 202 ACGAAATCGTTCTGTAGAAAAATGATTCGTCAATTAATCGAGTTTCAAGATTAGAGGTG 143
Qy 1337 CAACCAACATTGTTTCTT 1355
Db 142 AGACCAACATTCCCTCTCT 124

RESULT 4
LOCUS BF251052 637 bp mRNA linear EST 15-NOV-2001
DEFINITION BF251052 Coccidioides immitis spherule cDNA library Coccidioides
ACCESSION BF251052
VERSION BF251052.1 GI:16931195
KEYWORDS EST.
SOURCE Coccidioides immitis.
ORGANISM Coccidioides immitis.
REFERENCE 1 (bases 1 to 637)
AUTHORS Gardner M.J. and Kirkland T.
TITLE Generation of ESTs from Coccidioides immitis spherule cDNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Malcolm J. Gardner
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@igf.org.

FEATURES
source
    ..637
    /organism="Coccidioides immitis"
    /db_xref="taxon:5501"
    /clone="C1AAB85"
    /dev_stage="spherule"
    /lab_host="SOLR"
    /note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 123 a 261 c 163 g 90 t
ORIGIN
Query Match 6.3%; Score 220.2; DB 12; Length 637;
Best Local Similarity 60.4%; Pred. No. 1.5e-47;
Matches 380; Conservative 0; Mismatches 248; Indels 1; Gaps 1;

Qy 538 TACCCATCTTTGTAAGGACGATGGTGGCGGATGGTGGCGGACGCGGTATGCGCTTTGTTCT 597
Db 7 TTCCCATCATCATCAAGCGCCCTTCGGCGGGGTGGCGCGCATGCGTGTGTCGCGC 66
Qy 598 TCACCTGATGAGTCCGCAATTTGGACACAGACATCTCTGTAAGCTGAAGCGGCATTC 657
Db 67 GACGAGGAGTCCCTCCGGACTCTCTCGAGCGCGCCACCTCGGAGGCCAATCCGCTTC 126
Qy 658 GCGGACGGTTCGGTATATGTCGAGGCTGTGATTAAACCCAGACATGAGTGCAG 717
Db 127 GCACACGCAACCGTCTTCGTCGAGCGCTTCTTCACAGGCCCAAGACATCGAGTCCAG 186

```

```

Qy 718 ATCCTTCGCGATCGCACTGAGAGTTGTATACCTTTTATGAACGTGCTCACTGCAG 777
Db 187 CTGCTCGCGGACACCAACGCGCTCGTCCACTCTACGAGCGTGACTCGAGGTCAG 246
Qy 778 CGTGTCAACCAAAAGTTGTGAAATTCGCCAGCAGCATTTGGATCCAGAACTGCGT 837
Db 247 CGCGCCACCAAGAGTGTGAGCTCGCCCGGCAAGGACCTCCCGTCGAGCTCGC 306
Qy 838 GATCGAATTTGCGGATGACGATAAAGTTCTGCGCTCCATTTGGTTACAGGGGGCGGA 897
Db 307 GACAGCATCTCGCGCGAGCGCGTCAAGCTGGCAAGTCCGTACGTAACCGAACGCGCG 366
Qy 898 ACCGTGGAATTTCTGGTGCATGAAAGGGCAACCAACGCTTTTCATCGAAATGAACCCAGT 957
Db 367 ACCCGCGAGTTCTCGTGCAGCCAGCTCAACCGCTACTACTTCATCGAGATCAACCCCGC 426
Qy 958 ATCCAGGTTGAGCACACCGTGACTGAAAGAGTCAACCGAGGTGGACCTGGTGAAGCGCAG 1017
Db 427 ATCCAGGTCGAGCACACCATCACCGAGGATCACCGGATCGACATCGTCCGCGCCAG 486
Qy 1018 ATGCGTTGGCTGCTGGTGAACCTTGAAGGAATTTGGGTCTGACCCCAAGATAAGATCAAG 1077
Db 487 ATCCAGATCGCGCGCGCCACCTCGAGCAGCTCGCGCTCACCCAGGACCGCATCTCC 546
Qy 1078 ACCCAGGTGAGCAGCTGCGCGCATCACACGGAAGAT-CCAAACAACGGCTTCG 1136
Db 547 ACCAGAGCTTCGCATCCAGTGCAGATCACCAACCGAGATCCCAACAAAGGGGCTTCCA 606
Qy 1137 CCAGATACCGGAATCATCCGCGTACC 1165
Db 607 GCCCGACACCCGGTAAGATCGGAGTCTACC 635

RESULT 5
LOCUS A1109467 675 bp mRNA linear EST 19-APR-2001
DEFINITION GH08678.5prime GH Drosophila melanogaster head pot2 Drosophila
ACCESSION A1109467
VERSION A1109467.1 GI:3477791
KEYWORDS EST.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
REFERENCE 1 (bases 1 to 675)
AUTHORS Harvey D., Brokstein P., Hong L., Evans-Holm M., Su C., Tsang G.,
Lewis S., and Rubin G.M.
TITLE BDGP/HMI Drosophila EST Project
JOURNAL Unpublished (2001)
COMMENT Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Plate: 86 row: G column: 6
High quality sequence stop: 588.
Location/Qualifiers
    1..675
    /organism="Drosophila melanogaster"
    /db_xref="taxon:7227"
    /clone="GH08678"
    /sex="male and female"
    /dev_stage="adult"
    /lab_host="DH5 - alpha"
    /note="Organ: head; Vector: pOT2; Site 1: EcoRI; Site 2:
XhoI; Sized fractionated cDNAs were directly ligated into
pOT2. Plasmid cDNA library."
BASE COUNT 142 a 214 c 193 g 126 t
ORIGIN

```



```

Db 671 CATACCATGATACATCAGGGCAGGTTGTGGCAGCCATGTCGCTGTGTGCACAAAGCAGG 730
QY 2335 GCAGATGCTGTGA 2348
Db 731 GCTGATGCTGTGA 744

RESULT 7
BI618593
DEFINITION
  RH49459.5prime RH Drosophila melanogaster normalized Head pf1c-1
  Drosophila melanogaster cDNA clone RH49459 5 similar to
  BcdRNA:GH06348: FBAN0001516 GO: [pyruvate carboxylase (GO:0004736)]
  located on: 2R 46B14-46C1; 08/19/2001, mRNA sequence.
ACCESSION
  BI618593
VERSION
  BI618593.1 GI:15514118
KEYWORDS
  EST.
SOURCE
  fruit fly.
  Drosophila melanogaster
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
  1 (bases 1 to 640)
  Stapleton M., Brokstein P., Hong L., Tyler D., Berman B., Carlson
  J., Champe M., Chavez C., Dorsett V., Farfan D., Friese E., George
  R., Gonzalez M., Guarin H., Harris N., Li P., Liao G., Misra S.,
  Mungall C.J., Nunoo J., Pacheb J., Patagas V., Park S.,
  Phouanehavong S., Wan K., Yu C., Lewis S.S., Celisner S. and Rubin
  G.M.
  BDGP/HMM RH Drosophila EST Project
  Unpublished (2001)
  Contact: Stapleton, M.
  BDGP
  Lawrence Berkeley National Lab
  One Cyclotron Rd, Berkeley, CA 94720, USA
  Fax: 510 486 6798
  Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
  hit genomic AE003831: arm:2R [4876906.5137815]
  estimated-cyto:46B13-46E4: 08/19/2001
  Plate: RH.494 row: E column: 11
  High quality sequence stop: 543.
  Location/Qualifiers
    1..640
    /organism="Drosophila melanogaster"
    /db_xref="taxon:7227"
    /clone="RH49459"
    /clone_lib="RH Drosophila melanogaster normalized Head
    pf1c-1"
    /sex="male and female"
    /dev_stage="Adult"
    /lab_host="DH5-alpha Tona"
    /notes="Organ: head; Vector: pFLcl; Site: 1: XhoI; Site: 2:
    BamHI; Library was kindly generated by Piero Carninci at
    the RIKEN. The library was normalized and excised using
    Cre recombinase. Plasmid cDNA library."
  BASE COUNT 138 a 185 c 193 g 124 t
  ORIGIN
    Query Match 6.0%; Score 207.8; DB 13; Length 640;
    Best Local Similarity 59.1%; Pred. No. 3.2e-44;
    Matches 375; Conservative 0; Mismatches 257; Indels 3; Gaps 1;

QY 643 GCTGAAGCGGATTCGGCAGGGTTCGGTATGTCGAACGTGCTGATTAAACCCCGAG 702
Db 1 GCCAAGCGCGCTTGGCAATGGTGGCATGTTTCATTGAGAAGTTCATCGAGCGACCGGT 60
QY 703 CACATTGAAGTGCAGATCTTGGCGATCGCGATCGGAGAACTGTACACCTTTATGACGT 762
Db 61 CACATTGAGTTCACTGCTGGGAGACAGCGCCGGAAATGATGTGATCTGTACGACGT 120
QY 763 GACTGCTACATGAGCGTGTGCACAAAGTTGTGCAATTTGGCCGACGACGATTTG 822
Db 121 GACTGCTCCGTGCAGCGTGTGCCACAGAAAGTGTGGAGATCGCCCGACCGCCGCTG 180

```

```

QY 823 GATCCAGAACTCGCGTATCGCATTTGTGCGGATCAGTAAAGTTCTGCGCTCCATTGGT 882
Db 181 CCCATCGAAATCCGGGACAAGATGACGAGGACAGTGGCGCTTTGGCCGCCATGTGGGC 240
QY 883 TACCAGGGCGGGGAACCGTGGAAATCTTTGTGCGATGAAAAGGCAACACAGCTTTTCATC 942
Db 241 TATGAAACGCGGACACCGTGGAGTTCTTTGCGACGAGTCCGCGCACTTCTACTTCATC 300
QY 943 GAATGAACCCACGATGATCCAGTTGAGCACACCTGATGAGAGTACAGGAGTGGAC 1002
Db 301 GAAGTGAACCGCGCTCTACAAGTGGAGCACCGTACCGGAGGAGATACAGGATCGAT 360
QY 1003 CTGTGTAAGGCGCAGATCGCTTGGCTGCTGTGCAACCTTGAAGGAATTTGGGTCGACC 1062
Db 361 TTGTGTCAGTCGCAATCCGCGTGGCGAGGCGATGACCTTACCCGAATTTGGGTACACG 420
QY 1063 CAAGATGATGATCAAGACCCACGCTGCAGCACTGCAGTGCCTGCATCACCACGGAATCCCA 1122
Db 421 CAGGACAAAGATCGTCCCGCGTGGCTATGCCATTCACTGCGCTGTGACCCACGAAATCCG 480
QY 1123 AACACGCGCTTCCGCGCCAGATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCT 1182
Db 481 GCCAACGATTTTCAGCCCAACACCGCTGCGTGTGAGGTTCTTCGATCTGGCGAGGTTATG 540
QY 1183 GCGCTTCGCTGTGAGCGTG---CAGCTCAGCTCGTGTGCGAAATCACCAGCACACTTTGAC 1239
Db 541 GGCATTAGTTGGACAGTGCCTCGCTACCGGAGCCATCATTTGCCATACATACGAC 600
QY 1240 TCCATGCTGGTGAATGACCTGCGGTGTTCCGA 1274
Db 601 TCCCTGCTGGTCAAGGTGATCTCGCACGCCAGCGA 635

RESULT 8
BG910468
LOCUS
  602806142F1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4938203
  5', mRNA sequence.
ACCESSION
  BG910468
VERSION
  BG910468.1 GI:14290944
KEYWORDS
  EST.
SOURCE
  human.
  ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 715)
  NIH-MGC(http://mgi.nci.nih.gov/)
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapsb-remail.nih.gov
  Tissue Procurement: David N. Louis, M.D.
  cDNA Library Preparation: Life Technologies, Inc.
  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Incyte Genomics, Inc.
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: LLAM10873 row: d column: 12
  High quality sequence stop: 715.
  Location/Qualifiers
    1..715
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:4938203"
    /clone_lib="NCI_CGAP_Brn67"
    /tissue_type="anaplastic oligodendroglioma with 1p/19q
    loss"
    /lab_host="DH10B (T1 phage-resistant)"
    /notes="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI;
    Site: 2: SalI; Cloned unidirectionally. Primary: Oligo dt.
    Average insert size 2.3 kb. Constructed by Life

```

BASE COUNT 156 a 239 c 210 g 110 t

Technology. Note: this is a NCI_CGAP Library."

Query Match 5.9%; Score 205; DB 12; Length 715;
Best Local Similarity 56.6%; Pred. No. 1.9e-43;
Matches 401; Conservative 0; Mismatches 305; Indels 3; Gaps 1;

QY 725 GCGATCCGACTGGAGAGTTGTACACCTTTATGAACTGCTCTACTGTGAGCGTGGTC 784
Db 1 GGGACCAAGTATGGGAACATCCTGCACCTGTACGAGCGAGACTGCTCCATCCAGCGGGCG 60

QY 785 ACCAAAAAGTTGTCGAAATTCGCGCAGCACAGCATTTGGATCCAGAACTGGTGTATCGCA 844
Db 61 ACCAAGAGTGTGCGAGATTCGCCCGCGCCACCTGGACCCGAGCTTCGGAATCGGC 120

QY 845 TTTGTGCGGATGTCAGTAAAGTTCTGCCGCTCCATTGGTTTACAGGGCGCGGAACCGTGG 904
Db 121 TCACCAAGGACTCTGTGAAACTCGCTAAACAGGTGGGCTACGAGAACGCGAGCACCTGG 180

QY 905 AATTCCTTGGTCGATGAAAGGGGAAACACAGCTTTTCATCGAAATGAACCCAGTATCCAGG 964
Db 181 AGTTCTCTGGTGGACAGGACCGCAAGCACTACTTTCATCGAGGTCAACTCCCGCTGCAGG 240

QY 965 TTGAGCACACCGTACTGGAAGACTCACCGAGTGGACCTGTGTAAGCGCGAGATCGCT 1024
Db 241 TGGAGCACACCGTACAGAGAGATACCGACGCTGGTCCATGCTCAGATCCACG 300

QY 1025 TGGCTGTGTGCAACCTTGAAGGAATTTGGTCTGACCCCAAGATAAGATCAAGACCCACG 1084
Db 301 TGGCTGAGGCGAGAGCTACCCGACCTGGGCTTGGCGGAGGAGAACTCCGATCAAG 360

QY 1085 GTGAGCACTGCGAGTCCGATACACCGAAGATCAAAACACCGGTTCCGCCAGATA 1144
Db 361 GGTGTGCATCCAGTCCGGGTCAACACCGAGGACCCCGCGGAGCTTCAGCGCGACA 420

QY 1145 CCGGAATATCACCGCTACCGCTCACAGCGGAGCTGGGTTCTGTTGACGGTCCAG 1204
Db 421 CCGGCCATTTAGGTGTTCCGAGCGGAGAGGCGATCGGCTCGGATTAATGCTT 480

QY 1205 CTCAGCTC---GGTGGCGAAATCAACCGCACACTTTGACTCCATGCTGGTGAATAAGCT 1261
Db 481 CCGCTTCCAAGGAGCGCTCATCTGCGCCCACTACGACTCCCTGCTCAAGTCAATG 540

QY 1262 GCGTGTTCGACTTTGAACTGCTGTGCTCGTCACAGCGCGCTGGCTGAGTTCA 1321
Db 541 CCCACGCAAAAGACCAACCCACCGCGCCCAAGATGAGCAGGCGCTTTGCGAGTTCC 600

QY 1322 CCGTGTCTGTGTGCAACCAATTTGGTTCTTGTGCTGGTGTGCTCGGGAAGAGACT 1381
Db 601 GCGTCCGAGGTGGAAGACCAATCGCTTCTGTCGAGAATGTGCTCAACCAACAGCAGT 660

QY 1382 TCAGTTTCAAGCGCATCGCCACCGGATTTATCGGCGATCACCCACACT 1430
Db 661 TCCTGGCAGCACTGTGGACACCAAGTTTCATCGAGCGGAGAACCCAGACT 709

RESULT 9
BJ063722 641 bp mRNA linear EST 10-DBC-2001
LOCUS BJ063722 NIBB Mochii normalized Xenopus tailbud library
DEFINITION laevis cdna clone XL077c10 5', mRNA sequence.
ACCESSION BJ063722
VERSION BJ063722.1 GI:17470882
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 641)
REFERENCE Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
Y.

TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
source
1. .641
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone_lib="XL077c10"
/clone_lib="NIBB Mochii normalized Xenopus tailbud
library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
were oligo-dr primed and directionally cloned. Staging
according to Nieuwkoop and Faber. Library is subtracted
and was constructed by N. Garret and A.M. Zorn,
(Wellcome/CRC Institute)."

BASE COUNT 190 a 142 c 161 g 148 t
ORIGIN

Query Match 5.8%; Score 201.2; DB 13; Length 641;
Best Local Similarity 58.6%; Pred. No. 1.9e-42;
Matches 369; Conservative 0; Mismatches 258; Indels 3; Gaps 1;

QY 694 AACCCCGACACATTTGAATGTCAGATCTTGGCGATCGCACTGGAGAGTTGTACACTT 753
Db 5 AAAACCTCGCCACATTTGAATTCAGATCTCGGTGATAAATATGGAATGTGATTTG 64

QY 754 TATGAAGCTGACTGCTCAGTGCAGCGTCTGTCACCAAAAGTTGTCGAAATTCGCCAGCA 813
Db 65 TATGAAGAGACTGTTCCGTCAGAGAGGCAACAGAGGTGGTTGAAATAGCCCCAGCT 124

QY 814 CAGCATTTGGATCCAGAACTGCGTGTGATCGCATTTGTGCGATGCACTAAAGTTCTGCCGC 873
Db 125 GCACAGCTGGACCTTCAGCTCAGAGACAGACTGACCCATGACTCTGTAAACTGGCAAA 184

QY 874 TCATTTGGTTACAGGGCGGGAAACCGTGGAAATTTCTGTCATGAAAGGGCAACAC 933
Db 185 CAGTGGGTTATGAGAATGCCGCACTGTGGAGTTCTTGTGGATAAACATGGAAGCAT 244

QY 934 GTTTTCATCGAATGACCCACCTATCCAGTTGAGCACACCGTGTGAGAGTCAAC 993
Db 245 TACTTCATGAAGTCAACTCCAGACTCCAAAGTTGAGCATACAGTAACGGAAGAAATACA 304

QY 994 GAGTGGACCTGTGAAAGCGCAGATGCGCTTGGCTGCTGCTGCAACCTTGAAGGAATTG 1053
Db 305 GATGTTGACCTAGTCCATGCCAGATCCATGTTGAGAGGGAAGAGCTTCTTGACTT 364

QY 1054 GGTCTGACCAAGATGAAGATCAAGACCCAGCTGTCAGCACTGTCAGTGGCCGATCACACG 1113
Db 365 GGCCTCAAGCAAGAGAACATCCGTATCAATGGTTGTGCCATTCACTGTCAGAGTCACAACA 424

QY 1114 GAGATTCACAAACAAAGCTTCCGCCAGATACCGGAACCTATCACCGGTACCGTCACTA 1173
Db 425 GAAGACCTTTCAGTGGGTTTCAGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 484

QY 1174 GCGGAGCTGGCGTTCGTCTTGACGGTG---CAGCTCAGCTCGGTGGCGAAATCACCGCA 1230
Db 485 GAAGGAATGGGAATCCGACTAGACAGTGTCTGAGCTTTTCAAGAGCTGTCATTTCCCA 544

QY 1231 CACTTTGACTCCATGCTGGTGAATAATGAATGCCGCTGGTTCGAGCTTTGAACTGTGTT 1290
Db 545 CATTATGACTCCCTTCTGTTCAAGGTTATTGCAACCGGAAGAACCATCAAGTTGCCGCC 604

QY 1291 GCTCGTCACAGCGCGCTTGGCTGAGTTTC 1320
Db 605 ACAAAATGAGCAGAGCTTTGGCAGAGTTTC 634


```
RESULT 10
BF781506 887 bp mRNA linear EST 12-JAN-2001
LOCUS 602104410F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4222419
DEFINITION 5', mRNA sequence.
ACCESSION BF781506
VERSION BF781506.1 GI:12086542
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS 1 (bases 1 to 887)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue procurement: Jeffrey E. Green, M.D.
CDNA Library prepared by: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Inyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLN9809 row: d column: 04
High quality sequence stop: 752.
Location/Qualifiers
1..887
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4222419"
/lab_host="NCI CGAP Kid14"
/note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: Not I;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library. |"
BASE COUNT 188 a 282 c 249 g 168 t
ORIGIN
Query Match 5.7%; Score 197.4; DB 12; Length 887;
Best Local Similarity 56.0%; Pred. No. 2.4e-41;
Matches 394; Conservative 0; Mismatches 306; Indels 3; Gaps 1;
QY 746 TACACCTTTATGACGAGTCTGCTCACTGACGCTGCTCACCACCAAAAGTTGTGAAATTG 805
DB 6 TGCACCTGTACGACGAGATGCTCCATCCAGCTGGCACCAGAGGTGTAGATCG 65
QY 806 GGCAGCACAGCATTTGGATCCAGAACTGGTGATCGCATTTGTGCGGATGCAGTAAAGT 865
DB 66 CCCGTGTACCCACTGGATCCCAACTTCGCTCAGCTCTCACAGTGAATCTGTCAAAC 125
QY 866 TCTGCCGCTCCATTTGGTTACAGGCGCGGAAACCGTGGAAATCTTTGGTGCATGAAAGG 925
DB 126 TTGCCAAGCAGGTAGGCTATGAGAACCGCGGCACTGTGGAGTTCTCTGGTGGACACG 185
QY 926 GCAACACGTTTTTCATCGAAATGAACCCACGATATCCAGGTTGAGCACACCGTGAAG 985
DB 186 GCAAGCACTACTTCATCGAGGTCAAATTCGCGCTGCGAGTGGAGCACAGGTCAACG 245
QY 986 AAGTCACCGAGGTGGACCTGGTGAAGGCGGAGATGGCTTGGCTGTGGTGCACACCTGA 1045
DB 246 AGATCACAGATGTGGACCTGGTCCATGCTCAGATCCAGTGTCCGAAGGCGGAGGCTGC 305
QY 1046 AGGAATTTGGTCTTGACCAAGATAAGATCAAGACCCACGCTGAGCACTGCAGTGCAGCA 1105
DB 306 CTGACCTGGGCTTGGCGCAGGAGAACATCCGATCAATGCTGTGCAATTCAGTGTGGG 365
QY 1106 TCACCACGGAAGATCCAAACAACGGCTTCGCCCCAGATACCGGAACATATCACCGCGTACC 1165

Db 366 TCACCACGAGGACCTCTGCACGCGAGCTTCAGCCAGACACCGGCCGCAATTGAGTTTTCC 425
QY 1166 GCTCACACGAGCGGAGCTGGCGTTCGTCTGACGGTGCAGCTCAGCTCGTGGCGGAATCA 1225
DB 426 GGAGTGGTGAAGGATGGGATCCGCTGGACACGCTCTGCATTCAGGGCGCTGTCA 485
QY 1226 ---CGCACACTTTGACTCCATGCTGTGTAATGAAATGACCTGCGTGGTTCGACTTGA 1282
DB 486 TATCGCCCCACTATGACTCTCTGCTGCTGAGGTCATTCACACGCAAGAACACCCCA 545
QY 1283 CTGCTGTGCTCTGTCAGACGCGCGTGTGGCTGAGTTACCGTGTCTGGTGTTCGAACCA 1342
DB 546 CAGCTGCCACCAAGATGAGCAGAGCCCTCGGCCGAGTTCGCTGTCGAGGTGTAAGACCA 605
QY 1343 ACATTGTTTCTTCTCGTGGTGTGCTGGGGAAGAGGACTTCACTTCCAAAGCGCATCGCA 1402
DB 606 ACATCCCTTCTCTCAGAAATGTTCTCAACAACACGACGAGTTCCTGGCAGGCACATGG 665
QY 1403 CCGGATTTATCGCGGATCACCCACACCTCTCTCAGGCTCCACCC 1445
DB 666 CCCAGTTTCATGATGAGAACCTCTGAGCTGTTCCAGCTTCGGCC 708

RESULT 11
AU202878 515 bp mRNA linear EST 17-JUL-2001
LOCUS AU202878 unpublished oligo-capped cDNA library, stage L2
DEFINITION Caenorhabditis elegans cDNA clone yk813901 5', mRNA sequence.
ACCESSION AU202878
VERSION AU202878.1 GI:14933043
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 515)
AUTHORS Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.
and Sugano, S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished (2001)
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
1..515
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk813901"
/clone_lib="unpublished oligo-capped cDNA library, stage
L2"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L2"
/note="The AD-wrmcDNA library was generated with poly(A)+
RNA isolated from both hermaphrodite and male N2 worms of
all larval stages, embryos, adults and dauers and the
subsequent generation of cDNAs by poly(A) priming. The
cDNAs were cloned into pPC86"
BASE COUNT 146 a 103 c 124 g 142 t
ORIGIN
Query Match 5.7%; Score 196.6; DB 9; Length 515;
Best Local Similarity 62.9%; Pred. No. 2.7e-41;
Matches 304; Conservative 0; Mismatches 179; Indels 0; Gaps 0;
QY 640 GAAGCTGAAGCGGATTCGGCGAGGTTCCGTATATGTCGAACGCTGTGATTAAACCCC 699
```

Db 32 GAAGCTCAAGCTCGTTTGGAGCGGAAGCTCTTTTCGTTGAAAAGTTTGTGAGAGACCA 91
Qy 700 CAGCACATTTGAAGTGCAGATCTTTGGCGATCGCACTGGAGAGTTGTACACCTTTATGAA 759
Db 92 AGACATATTTGAAGTTTCAAGTCTTTGGAGACCATCATGGAATAATTTGTTCAATTTGTATGAG 151
Qy 760 CGTGACTGCTCACTGCGAGCTGCTCACCAAAAAGTTTGTGAAAATTTGGCCAGCAGCAT 819
Db 152 CGTGATTTGTTCAAGTGCAGCGCTCATCAAAAGTTTGTGAAAATTTGTCTCCAGCGCCAGCT 211
Qy 820 TTGGATCCAGAACTGCGTGAATGCGATGTCGAGTGCAGTAAAGTTTGTGCGCTCCATT 879
Db 212 CTCCAGAAAGTGTGCTGAGAAAATTTTGGCAGACGCTTCTCAGCTTGCAGACATGTT 271
Qy 880 GGTTCACGAGCGCGGGAACCGTGGAAATTTTGTGTCGATGAAAAGGGCAACACCGTTTC 939
Db 272 GGATACCAAAATGCTGCTGACAGTCAATCTCTGCTGATCAGAAAGGGCAACTACTATTTC 331
Qy 940 ATCGAATGACACCATTCACGTTTGGACACACCGTGAAGTGAAGTCAACGAGTG 999
Db 332 ATCGAAGTGAATGACGCTTCAAGTGCAGATACAGTAAGTGAAGATCACTGGTGTG 391
Qy 1000 GACCTGTGAAGGCGGAGATGCGCTTGGCTGCTGCTGCAACCTTTGAAGGAATTTGGGTCTG 1059
Db 392 GATCTTGTCCAGCTCAATTCGTATGCCGAAGAAAATCTCTGGATGATCTGAAGCTT 451
Qy 1060 ACCAAGATGAATCAAGACCCACCGTGCAGCTGCGCATCCACACGGAAGAT 1119
Db 452 TCACAGGAACATAATTCAAACTACTGGCTGAGCTATTCAATGTGCTGCACAACTGAAGAT 511
Qy 1120 CCA 1122
Db 512 CCA 514

RESULT 12
BO752438
LOCUS BO752438.1 DSCT Colletotrichum trifolii cDNA clone pDSCT12-67, mRNA
DEFINITION BO752438.1 GI:21907843
ACCESSION BO752438
VERSION EST.
KEYWORDS Colletotrichum trifolii.
SOURCE Colletotrichum trifolii.
ORGANISM Colletotrichum trifolii.
REFERENCE 1 (bases 1 to 949)
AUTHORS Samad,D.A., Dickman,M., Town,C.D., Van Aken,S., Utterback,T.,
Cheung,F. and Fraser,C.M.
TITLE ESTs from mycelia of Colletotrichum trifolii race 1
JOURNAL Unpublished (2002)
COMMENT Other_ESTs: EST633000
Contact: Deborah A. Samad
Department of Plant Pathology
University of Minnesota
495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
Tel: 612 625 1243
Fax: 612 625 5058
Email: debbys@pucini.crl.umn.edu
TIGR sequence name: MTSAL67TV More information is available at:
www.medicago.org
Seq primer: (gTA AtA CgA CtC Act ALC 99g C).
FEATURES
source Location/Qualifiers
1..949
/organism="Colletotrichum trifolii"
/strain="race 1"
/db xref="taxon:5466"
/clone="pDSCT12-67"
/clone_lib="DSCt"
/tissue_type="mycelia"
/dev_stage="Young, actively growing mycelia (3 days after

inoculation) grown in liquid culture (cutin minimal medium containing 2%glucose)."
/lab_host="DH5alpha"
/note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2sp2 ; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gt10 from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation and ligated into transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."
BASE COUNT 188 a 312 c 241 g 208 t
ORIGIN
Query Match 5.6%; Score 196.2; DB 14; Length 949;
Best Local Similarity 54.7%; Pred. No. 5.2e-41;
Matches 412; Conservative 0; Mismatches 338; Indels 3; Gaps 1;
Qy 2122 ACCCTGGATTACTACCTAAAGATGGCAGAGATCGTCAAGTCTGGCGCTCAATTCTG 2181
Db 1 AACCTCCAGTACTACCTGGATCTGGTGCAGAGCTGTTGCCCTCGACATCCAGTTCTC 60
Qy 2182 GCCATTAAAGATATGGTGTCTGTTGCCAGCTGCGGTAAACCAAGCTGGTCAACGCA 2241
Db 61 GGATCAAGGACATGGCGGTCTCTCAAGCCACCGCTACCTCTTTGATCGAGCC 120
Qy 2242 CTGGCGGTCAATTC---GATCTGCCAGTGCAGTGCACACCCACGACACTGCGGTGGC 2298
Db 121 ATCCGCAAGAAAGTACCTGACCTCCCTATTTCAGTTTCATCTACGATTCTGCTGGTACG 180
Qy 2299 CAGTTGGCTACTACTTTGCTGCAGCTCAAGCTGGTGGCAGATGCTGTGACGGTCTTC 2358
Db 181 GGTGTTGCTTCTATGTTGTTGGCCAGCGCGGTGCCAGCGCTGGAGCTGCCACC 240
Qy 2359 GCACCACTGTCTGGCACCACTCCAGCCATCCCTGCTCTGCCAATTTGTTGTGCAATTCGG 2418
Db 241 GACAGCTGTGTCGGCATGACCTCCAGCCAGCATCAAGCCATAATTTGTTCTCTCGAA 300
Qy 2419 CACACCGTGGCATACCGGTTTGAGCTCGAGCTGTTTCTGACCTCGAGCGCTACTGG 2478
Db 301 GGAGTGAACACGACACCGCGCTGAAACCTTGCAGCCCTGGAACACCTATTGG 360
Qy 2479 GAAGCTGTGCGCGGAGCTGTACCTGCCATTTGAGTCTGGAACCCAGGCCCAACCGGTGCG 2538
Db 361 TCTCAGCTGGCTTCTTCTACTCACCTTTGAGGCTCATCTCGTGTGCTCTGACCCCTGAG 420
Qy 2539 GTCTACCGCCACGAAATCCAGGCGGACAGTTGTCTCAACCTGCGTGCACAGGCCACCGCA 2598
Db 421 GTGTACGAGCATGAGATCCCGGTGGCCAGCTTTACCAACATGATGTTTCCAAGCAGCTCAG 480
Qy 2599 CTGGGCTTGTCTGATCGTCTGAGCTCATCGAAGACAACTACGAGCGCTTTATGAGATG 2658
Db 481 CTTGGTCTTTGGAACAACAATGGGCGGAGACCAAGAAAGCGTACGAGCATGCCAACGACCTC 540
Qy 2659 CTGGAGCGCCAAACCAAGGTCAACCTCTCCAAAGTGTGTGGCGACCTCGCACTCCAC 2718
Db 541 CTTGGCGACATTTGTCAGGTCATCTCCACTTCAAGTCTGTTGGCGACTTTGGCCAGTTTC 600
Qy 2719 CTGGTTGTGCGGGTGTAGATCCAGCAGACTTTTGTCTGCAGACCCCAAAAGTACGACATC 2778
Db 601 ATGTTTCGAACAAGCTTACCTCCCTGAGGATGTCAAGGCGCGCGCGGAGCTCGACTTC 660
Qy 2779 CAGACTCTGTCTATCGGTTCTTCGCGCGGAGCTTGGTAAACCTTCAGGTGGTGGCGCA 2838
Db 661 CCGGCTCTGTCTCGAGTTCTTGAAGGTTATGATGGGCGAGCAATTTGTGGTATTCGCC 720
Qy 2839 GAACCACTGCGCAACCGCGCAGCTGGAAGCGCGC 2871
Db 721 GAGCCCCCTTCGTCGAACGCTCTCCGTGGCGCGC 753

```

RESULT 13
BI605019
LOCUS
DEFINITION
  RH70532, Sprime RH Drosophila melanogaster normalized Head pFLC-1
  Drosophila melanogaster cDNA clone RH70532 5 similar to
  pCDNA:GH06348; Fban001516 GO: lpyruvate carboxylase (GO:0004736))
  located on: 2R 46B14-46C1; 08/24/2001, mRNA sequence.
ACCESSION
  BI605019
VERSION
  BI605019.1 GI:15500544
SOURCE
  EST.
  fruit fly.
ORGANISM
  Drosophila melanogaster
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
  1 (bases 1 to 666)
REFERENCE
  Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson
  J., Champagne, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George
  R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Misra, S.,
  Mungall, C. J., Nuncio, J., Pacleb, J., Paragas, V., Park, S.,
  Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Celniker, S. and Rubin
  G. M.
  BDGP/HMI RH Drosophila EST Project
  Unpublished (2001)
  Contact: Stapleton, M.
  BDGP
  Lawrence Berkeley National Lab
  One Cyclotron Rd, Berkeley, CA 94720, USA
  Fax: 510 486 6798
  Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
  hit genomic AE003831: arm:2R [4876906,5137815]
  estimated-cyto:46B13-46B4: 08/24/2001
  Plate: RH.705 row: C column: 8
  High quality sequence stop: 521.
  Location/Qualifiers
    1..666
      /organism="Drosophila melanogaster"
      /db_xref="taxon:7227"
      /clone="RH70532"
      /clone_lib="RH Drosophila melanogaster normalized Head
      pFLC-1"
      /sex="male and female"
      /dev_stage="Adult"
      /lab_host="DH5-alpha Tona"
      /note="Organ: head; Vector: pFLC1; Site 1: XhoI; Site 2:
      BamHI; Library was kindly generated by Piero Carninci at
      the RIKEN. The library was normalized and excised using
      Cre recombinase. Plasmid cDNA library."
  BASE COUNT      142 a 207 c 195 g 121 t
  ORIGIN

Query Match      5.8%; Score 195; DB 13; Length 666;
Best Local Similarity 58.4%; Pred. NO. 8.7e-41;
Matches 361; Conservative 0; Mismatches 256; Indels 3; Gaps 1;

QY 1561 AAGGATCGCACTGCCACGCGGTCCCGTGACCGCGTGAAGCAGCTTGGCCGCGCG 1620
DB 46 AAAGTTACGAACACCAACGGAAGGACTCCGCGAGGTCTCGTGTGCGAAGTCCGAGGCC 105
QY 1621 TTTCCTCGTATCTCCGTGAGCAGGACGCACTGGCGATTTACTGATACCACTTCGCGGAT 1680
DB 106 TTCGCCAAGGAGGTGCGCAACCGTAAGAACTGTCTACTATGAGACACCACTTCGCGGAT 165
QY 1681 GCACACCAAGTCTTTGCTTGCGACCCGAGTCCGCTCACTTCGCACTGAAGCCTGCGGAGAG 1740
DB 166 GCCCACCAGTCCGCTGTCGCGCACCCGCGGTGCTTCCACGATCTGCTGAAGATCTCCCCC 225
QY 1741 GCCGTCGCAAGCTGACTCTGAGCTTTTCTCGGTGAGGCTTGGCGCGCGGACCTAC 1800
DB 226 TAGGTGACGCACAAAGTTCAACAACTGTTATTCGCTGAGAACTGGGCGGAGGCACCTTC 285
QY 1801 GATGCGCGATCGGTTCTTCTTGTAGGATCCGTGGGACAGGCTCGACGAGCTGCGGAG 1860

```

```

DB 286 GACGTGGCGCTCGCTTCTCTGCACGAGTCCCGTGGGAGCGCTGGAGGAGATGCGGAG 345
QY 1861 GCGATGCGGATGTAAACATTTAGATGCTCTTCGCGGCGCAACACCGTGGGATACACC 1920
DB 346 CGATCCGAAACATTCCTTCCAGATGCTTTGGCGGAGCAACGCGTGGCTACACC 405
QY 1921 CCGTACCAGATCCGCTCGCGCGCTTTGTAAGCAAGCTGCCAGCTCGGCGTGCAC 1980
DB 406 AGCTATCCGGACACGCTGCTTACAGTTTGTGAGCTGCTGTGACAGCGGATGAC 465
QY 1981 ATCTTCGATCTTCGACGGCTTAAACGAGTCTCCAGATGCGTCCAGCAATCGACGA 2040
DB 466 ATCTTCAGGCTGTTCGACTCGCTCACTACCTGCCAACCTGATCTTCGCGATGAAGCC 525
QY 2041 GTCCTGGAGACCAACACCGCGGTAGCGAGTGGCTATGCTTATTCTGTGATCTCTCT 2100
DB 526 TGCTGAAGGCGCGCGCGTGGT---GGAGGCTGCCATCTNCTATACCGGAGAGTCA 582
QY 2101 GATCCAAATGAAAGCTCTTACACCTCGATTAACCTAAAGATGGCAGAGGAGATCGTC 2160
DB 583 GATCCCAAGCGCACCAAGTATGATCTGAAATACTACACTAACCTTGGCGGATGAGCTGTC 642
QY 2161 AAGTCTGCGCTCACATTCT 2180
DB 643 AAGGCGGCGCACCAAGTCT 662

RESULT 14
CNS06M6L
LOCUS
DEFINITION
  T3 end of clone AU00A002D08 of library AU00A from strain CBS 3082
  of Saccharomyces kluyveri, genomic survey sequence.
ACCESSION
  AL405091
VERSION
  AL405091.1 GI:12167396
KEYWORDS
  GSS.
SOURCE
  Saccharomyces kluyveri.
  Saccharomyces kluyveri.
REFERENCE
  Souciet, J. L., Aigle, M., Artiguenave, F., Blandin, G.,
  Bolotin-Fukuhara, M., Bon, E., Brottier, P., Casaregola, S.,
  de-Montigny, J., Dujon, B., Durrens, P., Lepingle, A., Llorente, B.,
  Malpertuy, A., Neuveglise, C., Ozier-Kalogeropoulos, O., Potier, S.,
  Saurin, W., Tekala, P., Toffano-Nioche, C., Wesolowski-Louvel, M.,
  Wincker, P. and Weissenbach, J.
  Genomic exploration of the hemiascomycetous yeasts: 1. A set of
  yeast species for molecular evolution studies
  FEBS Lett. 487 (1), 3-12 (2000)
  20584711
  PUBMED
  11152876
  2 (bases 1 to 1030)
  Neuveglise, C., Bon, E., Lepingle, A., Wincker, P., Artiguenave, F.,
  Gaillardin, C. and Casaregola, S.
  Genomic exploration of the hemiascomycetous yeasts: 9.
  Saccharomyces kluyveri
  FEBS Lett. 487 (1), 56-60 (2000)
  20584719
  PUBMED
  11152884
  3 (bases 1 to 1030)
  Genoscope.
  Direct Submission
  Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
  2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail:
  seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
  This GSS is part of a random genomic sequencing program of thirteen
  yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
  exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
  Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
  lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
  angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
  Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to

```

5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.

FEATURES

```

source          1. .1030
/organism="Saccharomyces kluveri"
/strain="CBS 3082"
/db_xref="taxon:4934"
/cloname="AU0AA002D08"
/clone_lib="AU0AA"
/notes="end : T3"
misc_feature    <12. .>1029
/note="similar to Saccharomyces cerevisiae ORF YBR218c [
PYC2 ; pyruvate carboxylase 2 ]
2 putative frameshift(s)
similar to Saccharomyces cerevisiae ORF YGL062w [ PYC1 ;
pyruvate carboxylase 1 ]"
/evidence=not_experimental
BASE COUNT      270 a      247 c      230 g      282 t
ORIGIN

```

Query Match	5.6%;	Score 193.2;	DB 17;	Length 1030;
Best Local Similarity	53.3%;	Pred. No. 3.5e-40;		
Matches	523;	Conservative 0;	Mismatches 423;	Indels 36; Gaps 4
Qy	892	GC	GGGAAACCGTGGAAATCTTTGGTCGATGA	AAAGGGCAACCGTTTTCATCGAAATGAAC 951
Db	21	GCCGGTACTCGCGAGTTTTTGGTCATCGCA	AAAAACAGACATATATTCATCGAAATCAAC 80	
Qy	952	CCAGGTATCCACAGGTTTGAGCACACCGTG	ACTGAAGAAGTCACCGAGGTGACCTGGTGAAG 1011	
Db	81	CCAAGATCCAAGTCAGACATACCATCTG	AGAAATACACCGGTGTCGACGTTGGGT 140	
Qy	1012	GCGCAGATGCGCTTGGCTGTGGTGCGAAC	CTTTGAAGGAATTTGGGTCTGACCAAGATAAG 1071	
Db	141	GCTCAAAATTCAAATTCGCCCGCGGCTT	CCTTGCAACAACATAGGCTTGTTCAGGATAAA 200	
Qy	1072	ATCAAGACCCACGGTGACGACTGACGATGC	CGCATCACACCGAAGATCCAACAACGGC 1131	
Db	201	ATAACCACTCGTGGTTGGCAATTCAAATG	TCGTATTAACCACTTGAAGATCTCTTCAAAGAAC 260	
Qy	1132	TTCCGCCCCAGATACCGGAACATCATACCG	GTACCGCTCACCAAGGGGAGCTGGCGTTGCT 1191	
Db	261	TTCCAACTGATACCGGTCGTATCGAAGTC	ATCTCGTTCCGCTGGTGTAAATGGTGTCA 320	
Qy	1192	CTTGACGGTG--CAGCTCAGCTCGGTGGG	GAATFACCGCACATTTTGACTTCCATGTG 1248	
Db	321	CTAGATGGTGTGAACCGTTTACGCGGTGT	GTCTCATTTCTCCACATTCGACTCTATGTTA 380	
Qy	1249	GTGAAATGACCTCGCGTGTTCGACTTTG	AACTGCTTTGAACTGCTGTTCGTCGCACACGGCGG 1308	
Db	381	GTCAGTGTTCCTGTTCCGGTCTCAACCT	ACGATACGAAATGTGTCGTAAGATGTTGGCGTGC 440	
Qy	1309	TTGGCTGAGTTACCGGTCTCTGGTTTGC	ACCAACATTTGGTTCTTGGCTGGTGTGCTG 1368	
Db	441	TTGATAGAATTCAGAAATCAGAGGTGT	TAAACCAACATTCCTTTCTACTTACTTGTCTA 500	
Qy	1369	CGGGAAGGAGACTTCATCTCCAGCGCAT	TCGCCACGGATTTATCGGCGATCACCCACAC 1428	
Db	501	ACCCATTTCCGTTTTCAGAGCGGTGAA	TATTGGACTACTTTTCATCGACACACTCCACAA 560	
Qy	1429	CTCCTTCAGGCTCCACCTGCGGATCAT	GAGCGGACGCATCCTGGATTACTTTCGAGAT 1488	
Db	561	CTATTTGAAATGGTTTCTTCTCAAAAC	AGAGCCCAAAAGTTGCTTACACTCTTGGCCGAC 620	
Qy	1489	GTCCCGGTGAACAAGCCTCATGGTGTG	CGGTCCAAAGGATGTTTGCAGCACCAATC---G 1543	
Db	621	TTGGCTGTCAATGGTTCTTCGATCA	AGGTCGTAATTTGGCCTACCAAAATTTGACCACTCAG 680	
Qy	1544	ATAAGTCCCAACATCAAGGATCTGCC	ACCGGTTCCCGGTGACCGC----- 1596	
Db	681	CCAAAGCTTCCAACTTTACAGCACT	CAAAACGGTAAACGTTTGTCCAGCTGTTGGCTCGCAC 740	

Qy	1597	-----CTGAAGCAGCTTGGCCAGCGCGCTTTGCTGCTGATCTCC	1636
Db	741	CAGCTGGCTGGAGAGATGTTTTTGTGGAGCAAGTCTCTGAAGGATTTTGCCAAACAAGTCC	800
Qy	1637	GTGAGCAGGACGCACTGGCAGTTACTGATACCACTTCCGCGATGCACACCAGTCTTTGC	1696
Db	801	GTAATAACAACGGTACTTTGTTGATGACACACCCTCGAGAGATGCCACCAATCTTTGC	860
Qy	1697	TTGGACCCGAGTCCGCTCATTTCCGACTGAAGCTCGGCGAGAGCGCGTGCAGAACTGA	1756
Db	861	TGGCTACCAGAGTTGTTACACACGATCTGGCTACTATTGCTCCAAACCACCTGCCCATGCTT	920
Qy	1757	CTCTCGAGCTTTTGTCCGTGGAGGCTGGGCGCGCGGACCTACGATGTGGCGATGCGTT	1816
Db	921	TGGCTGGTGTCTTTTGCCTTTTGGAAATGTTGGGTGTGTC-ACTTTCGAGCTTGCCATGAGAT	979
Qy	1817	TCCTCTTTGAGGATCCGTGGGA	1838
Db	980	TCCTACAGAAATCCATGGGA	1001
RESULT	15		
LOCUS	AI293242	625 bp	mRNA linear EST 19-APR-2001
DEFINITION	GH16378.5prime GH Drosophila melanogaster head pot2 Drosophila melanogaster cDNA clone GH16378 5prime, mRNA sequence.		
ACCESSION	AI293242		
VERSION	AI293242.1	GI:3942649	
KEYWORDS	EST.		
SOURCE	fruit fly.		
ORGANISM	Drosophila melanogaster		
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	1 (bases 1 to 625)		
AUTHORS	Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Teang, G., Lewis, S. and Rubin, G.M.		
TITLE	BDGP/HHMI Drosophila EST Project		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Stapleton, M. BDGP Lawrence Berkeley National Lab One Cyclotron Rd, Berkeley, CA 94720, USA Fax: 510 486 6798 Email: http://www.fruitfly.org/EST.._est@fruitfly.berkeley.edu Plate: 163 row: G column: 6 High quality sequence stop: 473.		
FEATURES	source		
	1..625		
	/location=Qualifiers		
	/organism="Drosophila melanogaster"		
	/db_xref="taxon:7227"		
	/clone="GH16378"		
	/clone_lib="GH Drosophila melanogaster head pot2"		
	/sex="male and female"		
	/dev_stage="adult"		
	/lab_host="DHS - alpha"		
	/note="Organ: head; Vector: pot2; Site 1: EcoRI; Site 2: XhoI; Sized fractionated cDNAs were directly ligated into pot2. Plasmid cDNA library."		
BASE COUNT	134 a 196 c 176 g 119 t		
ORIGIN			
Query Match	5.5%	Score 192.6;	DB 9; Length 625;
Best Local Similarity	60.2%	Pred. No. 3.7e-40;	
Matches 356;	Conservative	0; Mismatches 229;	Indels 6; Gaps 2;
Qy	1763	AGCTTTTGTCCGTGAGGCCTTGGGCGCGCGACCTACGATGTGGCGATGCGTTTCTCT	1822
Db	37	ACCTGTATTGCTGGAGAACTGGGCGGAGCCACCTTCGAGTGGCGCTGGCTTCTGCG	96
Qy	1823	TTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGGAGGCGATCCGAATGTAAACATTC	1882
Db	97	ACGAGTGCCTGGGAGCGCTGAGAGAGATGCCAAGCGCATCCGAAACATTCCCTTCC	156

```
QY 1883 AGATGCTGCTTCGGCGCCGCAACACCGTGGGATACACCCCGTACCAGACTCGTCTGCC 1942
Db |||||
QY 157 AGATGCTGTTGGCGGAGCCAAACGCGTGGCTACACAGCTATCCGGACAACGTGGTCT 216
Db |||||
QY 1943 GCGGTTTGTAAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGGCC 2002
Db |||||
QY 217 ACAAGTTCTGTGAGCTGGCTGTGCAGACCGGCATGGACATCTTCAGGGTCTTCGACTCGC 276
QY 2003 TTAACGAGCTCTCCAGATCGGTCCAGCATCGACCGAGTCCCTGGAGACCAACACCGCG 2062
Db |||||
QY 277 TCACTACCTCGCCCACTGATCTCTGGCATGGAGCCGCTGGAAGGGCGGGCGGTGG 336
QY 2063 TAGCCGAGGTGGCTATGGCTTATCTGGTGATCTCTGATCCAAATGAAAAGCTCTACA 2122
Db |||||
QY 337 T--GGAGGTGCCATCTCCTATACCGGAGACGTCAAGCGATCCCAAGCGCACCAAGTATG 393
QY 2123 CCTTGAATTACTACTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACTTCTGG 2182
Db |||||
QY 394 ATCTGAAATACTACACTAACCCTTGCCGATGAGCTGGTCAAGGGGGCACCCACGTGCTCT 453
QY 2183 CCATTAAAGGATATGGCTGTCTGCTTCCGCCAGCTGCGGTAAACCAAGCTGGTCAACGAC 2242
Db |||||
QY 454 GCATAAAGGATATGGCTGGTCTGCTCAAGCCCGAATCCGCCAGGCTTCTGATCACAGCCA 513
QY 2243 TGGC--CCGTGAATTGATCTGCCAGTGCACGTGCACACCCACGACACTGGCGGTGGCC 2299
Db |||||
QY 514 TTCGTGACAAGCACCCCGACATCCCAATCCATATCCACACCCACGACACCTCTGGAGCGG 573
QY 2300 AGTTGGCTACCTACTTTGCTGCGAGCTCAAGCTGGTGCAGATGCTGTGAGC 2350
Db |||||
QY 574 GAGTGGCCTTCATGCTGGCATGGCCCAATGCCGAGCCGATGTGGTGACG 624
```

Search completed: April 5, 2003, 07:14:38
Job time : 5257 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2003, 05:32:17 ; Search time 126 Seconds
(without alignments)
8455.516 Million cell updates/sec

Title: US-09-974-973-1
Perfect score: 3474
Sequence: 1 gtactgctatcacctctgg.....tgatcgctgctgttctctaa 3474

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3398.8	97.8	3621	4	US-09-220-081-1
2	3398.8	97.8	3621	4	US-09-677-575-1
3	476.6	13.7	3465	4	US-09-134-001C-591
4	306.6	8.8	1362	1	US-07-956-700B-5
5	306.6	8.8	1362	1	US-08-476-537-5
6	306.6	8.8	1362	1	US-08-485-607-5
7	306.6	8.8	1362	2	US-08-475-879-5
8	306.6	8.8	1362	4	US-09-433-043B-5
9	305	8.8	1362	1	US-08-611-107-7
10	305	8.8	1362	2	US-08-422-560A-7
11	305	8.8	1362	4	US-08-468-793-7
12	302.2	8.7	3077	1	US-08-074-121-1
13	302.2	8.7	3077	5	PCT-US94-06447-1
14	297	8.5	3065	1	US-07-956-700B-1
15	297	8.5	3065	1	US-08-476-537-1
16	297	8.5	3065	1	US-08-485-607-1
17	297	8.5	3065	1	US-08-611-107-5
18	297	8.5	3065	2	US-08-422-560A-5
19	297	8.5	3065	2	US-08-475-879-1
20	297	8.5	3065	4	US-08-468-793-5
21	297	8.5	3065	4	US-09-433-043B-1
22	267.4	7.7	1921	4	US-08-662-344-1
23	255	7.3	3231	1	US-08-074-121-4
24	255	7.3	3231	5	PCT-US94-06447-4
C 25	254.8	7.3	23673	4	US-09-773-816-1
C 26	247.4	7.1	19702	4	US-08-961-527-7
27	209	6.0	1236	4	US-09-134-001C-384

28	167.8	4.8	1386	4	US-09-134-001C-767	Sequence 767, App
29	136.8	3.9	1916	3	US-08-714-918-96	Sequence 96, Appl
30	136.8	3.9	1916	4	US-09-265-315-96	Sequence 96, Appl
31	136.8	3.9	1916	4	US-09-265-315-96	Sequence 96, Appl
32	136.8	3.9	1916	4	US-09-266-417-96	Sequence 96, Appl
C 33	108.2	3.1	4411529	4	US-09-103-840A-1	Sequence 1, Appli
34	99	2.8	4411529	4	US-09-103-840A-1	Sequence 1, Appli
C 35	88.2	2.5	762	4	US-08-998-416-387	Sequence 387, App
36	85.8	2.5	6270	1	US-08-418-893D-25	Sequence 25, Appl
37	85.8	2.5	6790	1	US-08-418-893D-22	Sequence 22, Appl
38	85	2.4	523	4	US-08-998-416-27	Sequence 27, Appl
39	68	2.0	7470	3	US-08-417-089-5	Sequence 5, Appli
40	68	2.0	7470	3	US-08-695-651-5	Sequence 5, Appli
41	68	2.0	7470	4	US-08-930-285-5	Sequence 5, Appli
42	68	2.0	7470	4	US-08-695-421-5	Sequence 5, Appli
43	68	2.0	7470	4	US-08-697-826A-9	Sequence 9, Appli
44	57.6	1.7	6993	4	US-08-934-386-6	Sequence 6, Appli
45	46.2	1.3	600	4	US-09-364-083-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-220-081-1
; Sequence 1, Application US/09220081
; Patent No. 6171833
; GENERAL INFORMATION:
; APPLICANT: Sinskey, Anthony J.
; APPLICANT: Lessard, Philip A.
; APPLICANT: Willis, Laura B.
; APPLICANT: Stephanopoulos, Gregory
; TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
; FILE REFERENCE: 1533.0790000
; CURRENT APPLICATION NUMBER: US/09/220,081
; CURRENT FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (199)..(3621)
US-09-220-081-1

Query Match	97.8%;	Score 3398.8;	DB 4;	Length 3621;
Best Local Similarity	98.6%;	Pred. No. 0;		
Matches 3427;	Conservative 0;	Mismatches 47;	Indels 0;	Gaps 0;
QY	1	GTGACTGCTATCACCCCTTGGCGGTCTCTTGTGAAAGGAATAATTACTCTAGTGTGCACT	60	
Db	148	GTGACTGCTATCACCCCTTGGCGGTCTCTTGTGAAAGGAATAATTACTCTAGTGTGCACT	207	
QY	61	CACACATCTTCAACGGCTTCAGCAATTCAAAAGATCTTGGTAGCAAAACCGCGCGGAATC	120	
Db	208	CACACATCTTCAACGGCTTCAGCAATTCAAAAGATCTTGGTAGCAAAACCGCGCGGAATC	267	
QY	121	CGCGTCCGTGCTTTCCGTGCGAGCACTCGAAACCGGTGACGCCACGGTAGCTATTATACCCC	180	
Db	268	CGCGTCCGTGCTTTCCGTGCGAGCACTCGAAACCGGTGACGCCACGGTAGCTATTATACCCC	327	
QY	181	CGTGAAGATCGGGGATCATTCACCGCTCTTTTGCTTCTTGAAGCTCTCGCATTTGGTACT	240	
Db	328	CGTGAAGATCGGGGATCATTCACCGCTCTTTTGCTTCTTGAAGCTCTCGCATTTGGTACC	387	
QY	241	GAAGGCTCACAGTCAAGCGGTACCTGGACATCGATGAAATTTATCGGTGCGAGCTAAAAA	300	
Db	388	GAAGGCTCACAGTCAAGCGGTACCTGGACATCGATGAAATTTATCGGTGCGAGCTAAAAA	447	
QY	301	GTTAAAGCAGATGCTATTTTACCGGGATATGGCTTCTCTGTGAAATATGCCAGCTTGCC	360	

Db 448 GTTAAAGCAGATGCCATTTACCCGGGATACGGCTTCTGTCTGAAATGCCAGCTGGCC 507
Qy 361 CGCAGTGC CGGAAAAACCGCATTTATTTATGGCCCAACCCACAGAGTTTCTTGATCTC 420
Db 508 CGCAGTGC CGGAAAAACCGCATTTATTTATGGCCCAACCCACAGAGTTTCTTGATCTC 567
Qy 421 ACCGGTGATAGTCTCGTGGGTAACCGCCGCGAAGAGCTGTCTGCCAGTTTGGCG 480
Db 568 ACCGGTGATAGTCTCGTGGGTAACCGCCGCGAAGAGCTGTCTGCCAGTTTGGCG 627
Qy 481 GAATCCACCCCGAGCAAAACATCGATGACATCGTTAAAGCCCTGAAGCCAGACTTAC 540
Db 628 GAATCCACCCCGAGCAAAACATCGATGACATCGTTAAAGCCCTGAAGCCAGACTTAC 687
Qy 541 CCGATCTTTGTAAGCAGATTGCCGTTGGTGGGAGCGCGGTATGCGCTTTGTTCTTCA 600
Db 688 CCGATCTTTGTAAGCAGATTGCCGTTGGTGGGAGCGCGGTATGCGCTTTGTTCTTCA 747
Qy 601 CCGATGAGCTCCGCAAAATTTGGCAACAGAGCATCTCGTGAAGCTGAAGCGGCAATTCGGC 660
Db 748 CCGATGAGCTTCGCAAAATTTAGCAACAGAGCATCTCGTGAAGCTGAAGCGGCTTCGGC 807
Qy 661 GACGGTTCCGTTATATGTCGAAACGTCGTGATTAACCCCGAGCAGCATTTGAAGTGCAGATC 720
Db 808 GATGGCGGTTATATGTCGAAACGTCGTGATTAACCCCGAGCAGCATTTGAAGTGCAGATC 867
Qy 721 CTTGGCGATCGCATCGAGAGTTGTACACCTTTATGAACGTCGCTCACTCGACGCT 780
Db 868 CTTGGCGATCACATCGAGAGTTGTACACCTTTATGAACGTCGCTCACTCGACGCT 927
Qy 781 CGTCAACAAAAGTTGTCGAAATTTGCCAGCAGCATTTGGATCCAGAACTCGCTGAT 840
Db 928 CGTCAACAAAAGTTGTCGAAATTTGCCAGCAGCATTTGGATCCAGAACTCGCTGAT 987
Qy 841 CGCATTTGCGGATCAGTAAGTTCTGCGGCTCAGTTGTTACAGGCGCGGGAAC 900
Db 988 CGCATTTGCGGATCAGTAAGTTCTGCGGCTCAGTTGTTACAGGCGCGGGAAC 1047
Qy 901 GTGGAATCTTGGTGCATGAAGGGCAACCACTTTTCATCGAAATGAACCCACGTAATC 960
Db 1048 GTGGAATCTTGGTGCATGAAGGGCAACCACTTTTCATCGAAATGAACCCACGTAATC 1107
Qy 961 CAGGTTGAGCACACCGTGACTGAAGAGTCAACCGAGGTGACCTGGTGAAAGGCGCAGATG 1020
Db 1108 CAGGTTGAGCACACCGTGACTGAAGAGTCAACCGAGGTGACCTGGTGAAAGGCGCAGATG 1167
Qy 1021 CGCTTGGCTGCTGGTCAACCTTGAAGGAATTTGGTCTGACCCCAAGATAAGATCAAGACC 1080
Db 1168 CGCTTGGCTGCTGGTCAACCTTGAAGGAATTTGGTCTGACCCCAAGATAAGATCAAGACC 1227
Qy 1081 CACGTCGACGACATGCAAGTCCGATCACACGGAAGATCCAAACACGCTTCCGCCCA 1140
Db 1228 CACGTCGACGACATGCAAGTCCGATCACACGGAAGATCCAAACACGCTTCCGCCCA 1287
Qy 1141 GATACCGGAATATCACGGGTACCGCTCACAGCGAGCTGGCTTCTTGAGCGT 1200
Db 1288 GATACCGGAATATCACGGGTACCGCTCACAGCGAGCTGGCTTCTTGAGCGT 1347
Qy 1201 GCAGTCAGCTCGGTGGGAAATCACCGCAACATTTGACATCCATGCTGGTGAAGTGAAC 1260
Db 1348 GCAGTCAGCTCGGTGGGAAATCACCGCAACATTTGACATCCATGCTGGTGAAGTGAAC 1407
Qy 1261 TGGCGTGGTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCTGGCTGAGTTTC 1320
Db 1408 TGGCGTGGTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCTGGCTGAGTTTC 1467
Qy 1321 ACCGTGTCTGGTGTGCAACCAATTTGGTTTCTTTCGGTGGTTCGTCGGGGAAGAGGAC 1380
Db 1468 ACCGTGTCTGGTGTGCAACCAATTTGGTTTCTTTCGGTGGTTCGTCGGGGAAGAGGAC 1527
Qy 1381 TTCACTTCCAGCGCATCGCCACCGGAATTTATCGGCGATCACCAACCTCTTCAGGCT 1440
Db 1528 TTCACTTCCAGCGCATCGCCACCGGAATTTATTCGCGATCACCCGCACTCTCTTCAGGCT 1587

Qy 1441 CCACCTCGGATGATGAGCAGGAGCGCATCTCTGGATTACTTTGGCAGATGTACCCGTGAAC 1500
Db 1588 CCACCTCGTGTATGATGAGCAGGAGCGCATCTCTGGATTACTTTGGCAGATGTACCCGTGAAC 1647
Qy 1501 AAGCCTCATGTGTGTCGCTCCAAAGGATGTTTCAGCAGCAACAAATCGATAGCTGCCAAATC 1560
Db 1648 AAGCCTCATGTGTGTCGCTCCAAAGGATGTTTCAGCAGCTCTATTCGATAGCTGCCAAATC 1707
Qy 1561 AAGGATCTGCCACCTGCCACGCGGTTCCCGTGAACCGCTGAAAGCAGCTTTGGCCAGCGCG 1620
Db 1708 AAGGATCTGCCACCTGCCACGCGGTTCCCGTGAACCGCTGAAAGCAGCTTTGGCCAGCGCG 1767
Qy 1621 TTTGCTCGTGTACTCTCCGTGAGCAGGAGCACTGGCAGTTACTGATACCACTTCGCGGAT 1680
Db 1768 TTTGCTCGTGTACTCTCCGTGAGCAGGAGCACTGGCAGTTACTGATACCACTTCGCGGAT 1827
Qy 1681 GCACACAGTCTTTGCTTGGACCGGAGTCCGCTCATTTGACACTGAAGCCTTCGCGCAG 1740
Db 1828 GCACACAGTCTTTGCTTGGACCGGAGTCCGCTCATTTGACACTGAAGCCTTCGCGCAG 1887
Qy 1741 GCGTTCGCAAAAGCTGACTCTGAGCTTTTGTCCGTGAGGCTGGGGCGCGCACTTAC 1800
Db 1888 GCGTTCGCAAAAGCTGACTCTGAGCTTTTGTCCGTGAGGCTGGGGCGCGCACTTAC 1947
Qy 1801 GATGTGGCGATCGCTTTCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTCGCGAG 1860
Db 1948 GATGTGGCGATCGCTTTCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTCGCGAG 2007
Qy 1861 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCCAACACCGTGGGATACACC 1920
Db 2008 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCCAACACCGTGGGATACACC 2067
Qy 1921 CCGTACCCAGACTCCGCTCGCGCGCTTCTTAAGGAAGCTGCAGCTCCGCGGTGGAC 1980
Db 2068 CCGTACCCAGACTCCGCTCGCGCGCTTCTTAAGGAAGCTGCAGCTCCGCGGTGGAC 2127
Qy 1981 ATCTTCGCGATCTTCGACGCGCTTAAACAGCTCTCCAGATGCGTCCAGCAATGACGCA 2040
Db 2128 ATCTTCGCGATCTTCGACGCGCTTAAACAGCTCTCCAGATGCGTCCAGCAATGACGCA 2187
Qy 2041 GTCTGGAGACCAACACCGCGGTAGCCAGTGGCTATGCTTATTTCTGGTGTCTCTCT 2100
Db 2188 GTCTGGAGACCAACACCGCGGTAGCCAGTGGCTATGCTTATTTCTGGTGTCTCTCT 2247
Qy 2101 GATCCAAATGAAAAGCTCTACACCTCGATTAACCTTAAAGATGGCAGAGAGATCGTC 2160
Db 2248 GATCCAAATGAAAAGCTCTACACCTCGATTAACCTTAAAGATGGCAGAGAGATCGTC 2307
Qy 2161 AAGTCTGGCGCTCACTTCTGGCCATTAAAGGATATGGCTGTCTTCCGCCAGCTGCG 2220
Db 2308 AAGTCTGGCGCTCACTTCTGGCCATTAAAGGATATGGCTGTCTTCCGCCAGCTGCG 2367
Qy 2221 GTAACCAAGCTGGTCAACCGGACTGGCGGTAATTCGATCTGCCAGTGCAGTGCACACC 2280
Db 2368 GTAACCAAGCTGGTCAACCGGACTGGCGGTAATTCGATCTGCCAGTGCAGTGCACACC 2427
Qy 2281 CACGACATCGCGGTGGCAGTGGCTTACCTTTCGTCAGCTCAAGCTGGTGCAGAT 2340
Db 2428 CACGACATCGCGGTGGCAGTGGCTTTCGTCAGCTCAAGCTGGTGCAGAT 2487
Qy 2341 GCTGTGTAGCGTGTCTCCGACCACTGTCTGGCAACCTCCAGGCCATTCCTGTCTGCC 2400
Db 2488 GCTGTGTAGCGTGTCTCCGACCACTGTCTGGCAACCTCCAGGCCATTCCTGTCTGCC 2547
Qy 2401 ATTGTGTGTGATTCGCGCACACCGCTCGCATACCGGTTTGGCCTCGAGGCTGTCTCT 2460
Db 2548 ATTGTGTGTGATTCGCGCACACCGCTCGCATACCGGTTTGGCCTCGAGGCTGTCTCT 2607
Qy 2461 GACCTCGAGCGGTACTCTGGGAAGCTGTGCGGAGCTGTACCTGCCATTTGAGTCTGGGAAC 2520
Db 2608 GACCTCGAGCGGTACTCTGGGAAGCTGTGCGGAGCTGTACCTGCCATTTGAGTCTGGGAAC 2667

```
QY 2521 CCAGGCCAACCCGGTCGGCTCTACCGCCACGAAATCCAGCGGACAGTGTCTCCAACTG 2580
Db 2568 CCAGGCCAACCCGGTCGGCTCTACCGCCACGAAATCCAGCGGACAGTGTCTCCAACTG 2727
QY 2581 CGTGACAGGCCACCGCACCTGGGCTTGTCTGATCGCTTCGAGCTCATCGAAGACAACTAC 2640
Db 2728 CGTGACAGGCCACCGCACCTGGGCTTGTCTGATCGCTTCGAGCTCATCGAAGACAACTAC 2787
QY 2641 GCAGCCGTTAATGATGCTGGAGCGGCCAACCAAGGTCAACCCATCTCCAAAGTTGTT 2700
Db 2788 GCAGCCGTTAATGATGCTGGAGCGGCCAACCAAGGTCAACCCATCTCCAAAGTTGTT 2847
QY 2701 GGGCACTCGCACTCCACCTGTTGGTGGGTGTAGATCCAGCAGACTTGTCTGCAGAC 2760
Db 2848 GGGCACTCGCACTCCACCTGTTGGTGGGTGTAGATCCAGCAGACTTGTCTGCAGAT 2907
QY 2761 CCACAAAGTACGACATCCCAAGCTCTGTATCGCTTCCTGCGCGGAGCTTTGGTAAC 2820
Db 2908 CCACAAAGTACGACATCCCAAGCTCTGTATCGCTTCCTGCGCGGAGCTTTGGTAAC 2967
QY 2821 CCTCAGGTGGCTGGCCAGAACCACTGGGACCCGGGCACTGGAAAGGCGCTCCGAAGGC 2880
Db 2968 CCTCAGGTGGCTGGCCAGAACCACTGGGACCCGGGCACTGGAAAGGCGCTCCGAAGGC 3027
QY 2881 AAGGCACCTCTGACGGAAGTTCTGAGGAAGAGCAGCGCACCTCGAGCTGATGATTC 2940
Db 3028 AAGGCACCTCTGACGGAAGTTCTGAGGAAGAGCAGCGCACCTCGAGCTGATGATTC 3087
QY 2941 AAGGAAGCTGCAACAGCCTCAACCGGCTGCTGTTCCGGAAGCCAAACGAAGAGTCTCTC 3000
Db 3088 AAGGAAGCTGCAACAGCCTCAACCGGCTGCTGTTCCGGAAGCCAAACGAAGAGTCTCTC 3147
QY 3001 GAGCACCGTCCCGCTTCGGCAACACCTCTGCGTGTGATGATGATGATGATGATGATG 3060
Db 3148 GAGCACCGTCCCGCTTCGGCAACACCTCTGCGTGTGATGATGATGATGATGATGATG 3207
QY 3061 CTGCTCGAGGCGCGGAGACTTTCATCGGCTGCGAGATGTCGACACCCACTGCTGTT 3120
Db 3208 CTGCTCGAGGCGCGGAGACTTTCATCGGCTGCGAGATGTCGACACCCACTGCTGTT 3267
QY 3121 CGCTTGATGCGATCTCTGAGCCAGACGATTAAGGCTATGCGCAATGTTGTGGCCAAAGTC 3180
Db 3268 CGCTTGATGCGATCTCTGAGCCAGACGATTAAGGCTATGCGCAATGTTGTGGCCAAAGTC 3327
QY 3181 AAGGCCAGATCCGCCCAATGCGTGTGCGTGAACCGCTCCGTTGATGCTGTCAACCGCAACC 3240
Db 3328 AAGGCCAGATCCGCCCAATGCGTGTGCGTGAACCGCTCCGTTGATGCTGTCAACCGCAACC 3387
QY 3241 GCAGAAAGGCGAGATTCCTCCAAAGGGCCATGTTGCTGACCACTTCGCTGCTGTTGTC 3300
Db 3388 GCAGAAAGGCGAGATTCCTCCAAAGGGCCATGTTGCTGACCACTTCGCTGCTGTTGTC 3447
QY 3301 ACTGTGACTGTTGCTGAAGGTGATGAGTCAAGGCTGAGATGAGTCAAGTCAATCATCGAG 3360
Db 3448 ACTGTGACTGTTGCTGAAGGTGATGAGTCAAGGCTGAGATGAGTCAAGTCAATCATCGAG 3507
QY 3361 GCTATGAAGTGAAGCAACAATCACTGCTTCTGTTGAACGCAAGATGGAACGGTTGTG 3420
Db 3508 GCTATGAAGTGAAGCAACAATCACTGCTTCTGTTGAACGCAAAATCGATCGGTTGTG 3567
QY 3421 GTTCTCTGCAACGAAGGTGAAGGTGGGCACTTGATCGTCTGCTGTTCTTAA 3474
Db 3568 GTTCTCTGCAACGAAGGTGAAGGTGGGCACTTGATCGTCTGCTGTTCTTAA 3621
```

RESULT 2

```
US-09-677-575-1
; Sequence 1, Application US/09677575
; Patent No. 6403351
; GENERAL INFORMATION:
; APPLICANT: Sinskey, Anthony J.
; APPLICANT: Lessard, Philip A.
; APPLICANT: Willis, Laura B.
```

```
; APPLICANT: Stephanopoulos, Gregory
; TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
; FILE REFERENCE: 1533.0790000
; CURRENT APPLICATION NUMBER: US/09/677,575
; CURRENT FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 09/220,081
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (199)..(3621)
US-09-677-575-1

Query Match 97.8%; Score 3398.8; DB 4; Length 3621;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 3427; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
```

```
QY 1 GTGACTCTATCACCCCTTGGCGGTCTCTTGTGAAAGAAATAATTACTCTAGTGTGCACT 60
Db 148 GTGACTCTATCACCCCTTGGCGGTCTCTTGTGAAAGAAATAATTACTCTAGTGTGCACT 207
QY 61 CACACATCTTCAACGGTTCAGCATTTCAAAAAGATCTTGTAGCAAAACCGCGGCAAAATC 120
Db 208 CACACATCTTCAACGGTTCAGCATTTCAAAAAGATCTTGTAGCAAAACCGCGGCAAAATC 267
QY 121 GCGGTCCGTCTTTCCTGTCAGCACTCGAAACCGGTGCAGCCAGGTAGCTATTATACCC 180
Db 268 GCGGTCCGTCTTTCCTGTCAGCACTCGAAACCGGTGCAGCCAGGTAGCTATTATACCC 327
QY 181 CGTGAAGATCGGGATCATTTCCACCGCTCTTTTCTCTCAAGCTGTCCGATTTGGTACT 240
Db 328 CGTGAAGATCGGGATCATTTCCACCGCTCTTTTCTCTCAAGCTGTCCGATTTGGTACT 387
QY 241 GAAGGTCCACAGTCAAGCGGTACCTGGACATCGATGAATAATTATCGGTGACGCTTAAAA 300
Db 388 GAAGGTCCACAGTCAAGCGGTACCTGGACATCGATGAATAATTATCGGTGACGCTTAAAA 447
QY 301 GTTAAAGCAGATGCTATTTACCCGGGATATGGCTTCTGTCTGAAATGCCCAGCTTGCC 360
Db 448 GTTAAAGCAGATGCTATTTACCCGGGATATGGCTTCTGTCTGAAATGCCCAGCTTGCC 507
QY 361 CGCGAGTGCSCGGAACCGCATTTACTTTTATTTGGCCCAACCCAGAGGTTCTTGATCTC 420
Db 508 CGCGAGTGCSCGGAACCGCATTTACTTTTATTTGGCCCAACCCAGAGGTTCTTGATCTC 567
QY 421 ACCGGTGATTAAGTCTCGTGCAGTAAACCGCGCGAAGAGGCTGTGCTGCCAGTTTGGCG 480
Db 568 ACCGGTGATTAAGTCTCGCGGGTAAACCGCGCGAAGAGGCTGTGCTGCCAGTTTGGCG 627
QY 481 GAATCAACCCCGAGCAAAAACATCGATGACATGTTAAAGCGCTGAAGCGCAGACTTAC 540
Db 628 GAATCAACCCCGAGCAAAAACATCGATGACATGTTAAAGCGCTGAAGCGCAGACTTAC 687
QY 541 CCATCTTTCTTAAGCAGTTTCCGCGTGGGCGGACGCGGTATCGCTTTGTTCTTCA 600
Db 688 CCATCTTTCTTAAGCAGTTTCCGCGTGGGCGGACGCGGTATCGCTTTGTTCTTCA 747
QY 601 CTTGATGAGTCCGCAAAATTTGGCAACAGAAAGCATCTCGTGAAGCTGAAGCGGCAATTCGGC 660
Db 748 CTTGATGAGTCCGCAAAATTTAGCAACAGAAAGCATCTCGTGAAGCTGAAGCGGCTTCGGC 807
QY 661 GACGTTTGGTATATGTCGAAACGTCGTGATTAACCCCGCAGCAATTTGAAGTGCAGATC 720
Db 808 GATGGCGCGGTATATGTCGAAACGTCGTGATTAACCCCGCAGCAATTTGAAGTGCAGATC 867
QY 721 CTTGGCGATCGCACTGGAGAGTTGTACACTTTATGAACGTCGCTCACTGCAGGT 780
Db 868 CTTGGCGATCACACTGGAGAGTTGTACACTTTATGAACGTCGCTCACTGCAGGT 927
```

Qy 781 CGTCAAAAAGTTGTCGAAATTCGCCAGACACAGATTTGGATCCAGAACTGGGTGAT 840
Db 928 CGTCAAAAAGTTGTCGAAATTCGCCAGACACAGATTTGGATCCAGAACTGGGTGAT 987
Qy 841 CGCATTTGCGGATGAGTAAGTTCTGCGGCTCCATTTGGTTACAGGGCGCGGAACC 900
Db 988 CGCATTTGCGGATGAGTAAGTTCTGCGGCTCCATTTGGTTACAGGGCGCGGAACC 1047
Qy 901 GTGGAATTTCTTGGTTCGATGAAAGGGCAACACCGTCTTTCATCGAAATGAACCCAGTATC 960
Db 1048 GTGGAATTTCTTGGTTCGATGAAAGGGCAACACCGTCTTTCATCGAAATGAACCCAGTATC 1107
Qy 961 CAGGTTGAGCACACCGTGAATCTGAAGAAAGTCAACGAGGTGACCTGGTGAAGGCGCAGATG 1020
Db 1108 CAGGTTGAGCACACCGTGAATCTGAAGAAAGTCAACGAGGTGACCTGGTGAAGGCGCAGATG 1167
Qy 1021 CGCTTGGCTGCTGGTGCACCTTTGAAGGAATTTGGGTCTGACCCCAAGATGAAGATCAAGACC 1080
Db 1168 CGCTTGGCTGCTGGTGCACCTTTGAAGGAATTTGGGTCTGACCCCAAGATGAAGATCAAGACC 1227
Qy 1081 CAGGTTGAGCACACCGTGAATCTGAAGAAAGTCAACGAGGTGACCTGGTGAAGGCGCAGATG 1140
Db 1228 CAGGTTGAGCACACCGTGAATCTGAAGAAAGTCAACGAGGTGACCTGGTGAAGGCGCAGATG 1287
Qy 1141 GATACCGGAATCTACCGGTGACCTGACCGGTGACCGGTGACCGGTGACCGGTGACCGGT 1200
Db 1288 GATACCGGAATCTACCGGTGACCTGACCGGTGACCGGTGACCGGTGACCGGTGACCGGT 1347
Qy 1201 GAGCTCAGCTCGGTGCGGAAATCAACGCACTTTGACTCCATGCTGCTGAAATGAGC 1260
Db 1348 GAGCTCAGCTCGGTGCGGAAATCAACGCACTTTGACTCCATGCTGCTGAAATGAGC 1407
Qy 1261 TGCGGTGTTCCGACTTTGAACTGCTGTTGCTCGTGACAGCGCGGTTGGTGAAGTTC 1320
Db 1408 TGCGGTGTTCCGACTTTGAACTGCTGTTGCTCGTGACAGCGCGGTTGGTGAAGTTC 1467
Qy 1321 ACCGTGCTGCTGTTGCAACCAATTTGTTCTGCGTGGTGGTGGCGGGAAGAGGAC 1380
Db 1468 ACCGTGCTGCTGTTGCAACCAATTTGTTCTGCGTGGTGGTGGCGGGAAGAGGAC 1527
Qy 1381 TTCACTTCCAAAGCGCATCGCCACCGGATTTATCGCGATCAACCACTCTTCAGGCT 1440
Db 1528 TTCACTTCCAAAGCGCATCGCCACCGGATTTATCGCGATCAACCACTCTTCAGGCT 1587
Qy 1441 CCACCTGCGGATGATGACGAGGACGATCTCGGATTAATTTGGCAGATGTCACCGGTGAC 1500
Db 1588 CCACCTGCGGATGATGACGAGGACGATCTCGGATTAATTTGGCAGATGTCACCGGTGAC 1647
Qy 1501 AAGCTCATGCTGCTGCTCAAGGATTTGCGACCAATCGATGAAGCTGCCCAACATC 1560
Db 1648 AAGCTCATGCTGCTGCTCAAGGATTTGCGAGCTCTATCGATGAAGCTGCCCAACATC 1707
Qy 1561 AAGGATCTGCACTGCCACCGGTTCCCGTGACCGCTGAAGCAGCTTGCGCCAGCGCG 1620
Db 1708 AAGGATCTGCACTGCCACCGGTTCCCGTGACCGCTGAAGCAGCTTGCGCCAGCGCG 1767
Qy 1621 TTTGCTGCTGATCTCCGTTGAGCAGGACGCACTGSCAGTTACTGATACCACTTCCGCGAT 1680
Db 1768 TTTGCTGCTGATCTCCGTTGAGCAGGACGCACTGSCAGTTACTGATACCACTTCCGCGAT 1827
Qy 1681 GCACACAGCTCTTTGCTTGGACCCGAGTCCGCTCATTTGCACTGAAGCCTTGCGGAGAG 1740
Db 1828 GCACACAGCTCTTTGCTTGGACCCGAGTCCGCTCATTTGCACTGAAGCCTTGCGGAGAG 1887
Qy 1741 GCGCTGCAAGAGCTGACTCTGCTGAGCTTTGTCGTTGAGGCTTGCGGCGCGGACCTAC 1800
Db 1888 GCGCTGCAAGAGCTGACTCTGCTGAGCTTTGTCGTTGAGGCTTGCGGCGCGGACCTAC 1947
Qy 1801 GATGTCGCGATGCGTTTCTTTGAGGATCCGTTGGGACAGGCTTCGACGAGCTGCGCGAG 1860
Db 1948 GATGTCGCGATGCGTTTCTTTGAGGATCCGTTGGGACAGGCTTCGACGAGCTGCGCGAG 2007

Qy 1861 CGGATCGCGAATGTAAACATTTAGATCTGCTTCCGCGCGCAACACCGTGGGATACACC 1920
Db 2008 CGGATCGCGAATGTAAACATTTAGATCTGCTTCCGCGCGCAACACCGTGGGATACACC 2067
Qy 1921 CCGTACCCAGACTCCGCTGCGCGCGCTTTGTTAAGAAAGTCCGAGCTCCGCGGTGGAC 1980
Db 2068 CCGTACCCAGACTCCGCTGCGCGCGCTTTGTTAAGAAAGTCCGAGCTCCGCGGTGGAC 2127
Qy 1981 ATCTTCCGATCTTTCGAGCGCGCTTAAACAGCTCTCCAGATGCGTCAGCAATCGAGCA 2040
Db 2128 ATCTTCCGATCTTTCGAGCGCGCTTAAACAGCTCTCCAGATGCGTCAGCAATCGAGCA 2187
Qy 2041 GTCTGGAGACAAACACCGCGGTAGCGAGTGGCTATGCGTATTTCTGTTGATCTCTCT 2100
Db 2188 GTCTGGAGACAAACACCGCGGTAGCGAGTGGCTATGCGTATTTCTGTTGATCTCTCT 2247
Qy 2101 GATCCAAATGAAAAGCTCTACACCGCTGATTAACCTACCTAAAGATGGCAGAGAGATGTC 2160
Db 2248 GATCCAAATGAAAAGCTCTACACCGCTGATTAACCTACCTAAAGATGGCAGAGAGATGTC 2307
Qy 2161 AAGTCTGGCGCTCAATTTCTGGCCATTAAGGATATGGCTGCTCTGCTTCCGCCAGCTGCG 2220
Db 2308 AAGTCTGGCGCTCAATTTCTGGCCATTAAGGATATGGCTGCTCTGCTTCCGCCAGCTGCG 2367
Qy 2221 GTTACCAAGCTGCTCACCGCACTCGCGGTGAATTCGATCTGCCAGTGCAGTGCACACC 2280
Db 2368 GTTACCAAGCTGCTCACCGCACTCGCGGTGAATTCGATCTGCCAGTGCAGTGCACACC 2427
Qy 2281 CACGACACTGCGGGTGGCGAGTTGGCTACCTTACCTTCTGCTGAGCTCAAGCTGGTGGAT 2340
Db 2428 CACGACACTGCGGGTGGCGAGTTGGCTACCTTACCTTCTGCTGAGCTCAAGCTGGTGGAT 2487
Qy 2341 GCTGTTGACGCTGCTTCCGCACTGCTGCGCACTGCTTCCAGCCATTCCTCTGCTGTC 2400
Db 2488 GCTGTTGACGCTGCTTCCGCACTGCTTCCGCACTGCTTCCAGCCATTCCTCTGCTGTC 2547
Qy 2401 ATTGTTGCTGATTTCCGCGCACACCGCTGCGATACCGGTTTGAAGCTCGAGGCTGTTCT 2460
Db 2548 ATTGTTGCTGATTTCCGCGCACACCGCTGCGATACCGGTTTGAAGCTCGAGGCTGTTCT 2607
Qy 2461 GACCTTCAGACCGCTACTGCGGAAAGCTGTCGCGGACTGTACCTTGCCTTGAAGTCTGGAACC 2520
Db 2608 GACCTTCAGACCGCTACTGCGGAAAGCTGTCGCGGACTGTACCTTGCCTTGAAGTCTGGAACC 2667
Qy 2521 CCAGGCCCCAAACCGGTCGCTCTACCGCCACGAAATCCAGCGCGACAGTTTGTCCAACTG 2580
Db 2668 CCAGGCCCCAAACCGGTCGCTCTACCGCCACGAAATCCAGCGCGACAGTTTGTCCAACTG 2727
Qy 2581 CGTGCACAGGCGCACCGCACTGGGCTTGTGATCGCTTCCAGCTCATCGAAGACAACTAC 2640
Db 2728 CGTGCACAGGCGCACCGCACTGGGCTTGTGATCGCTTCCAGCTCATCGAAGACAACTAC 2787
Qy 2641 CGAGCGGTTAATGATGCTGGGACCGCCCAACCAAGTCACTCCCATCTCTCAAGGTTGTT 2700
Db 2788 CGAGCGGTTAATGATGCTGGGACCGCCCAACCAAGTCACTCCCATCTCTCAAGGTTGTT 2847
Qy 2701 GCGCAGCTCGCACTCCACTGCTGCTGGTGGTGGGTTAGATCCAGCAGACTTTGCTGCGAGAC 2760
Db 2848 GCGCAGCTCGCACTCCACTGCTGCTGGTGGTGGGTTAGATCCAGCAGACTTTGCTGCGAGAC 2907
Qy 2761 CCACAAAAGTACGACATCCAGACTCTGTATCGGTTTCTGCGCGGAGCTTGGTAAAC 2820
Db 2908 CCACAAAAGTACGACATCCAGACTCTGTATCGGTTTCTGCGCGGAGCTTGGTAAAC 2967
Qy 2821 CCTCCAGGTGGCTGGCGCAGAAACCACTGCGCACCGCGCACTGGAAGGCGCTCCGGAAGGC 2880
Db 2968 CCTCCAGGTGGCTGGCGCAGAGCACTGCGCACCGCGCACTGGAAGGCGCTCCGGAAGGC 3027
Qy 2881 AAGGACCTCTGACGGAAGTTCTTGAGGAAGAGCAGGCGCACCTCGAGCTGATGATTC 2940
Db 3028 AAGGACCTCTGACGGAAGTTCTTGAGGAAGAGCAGGCGCACCTCGAGCTGATGATTC 3087
Qy 2941 AAGGACCTCTGACGCAACCGCTCTCAACCGCTGCTGTTCCGGAAGCCAAACCGAGATTCCTC 3000

Db 3088 AAGGAAGTGCATAGAGCTCAACCGCTGCTGTTCCGAAAGCCAAACGAAAGATGTTCTC 3147
Qy 3001 GAGCACCGTCCCGCTTCGGCAACACCTCTCGCTGCTGATGATCGTGAATTTCTTCAAGGA 3060
Db 3148 GAGCACCGTCCCGCTTCGGCAACACCTCTCGCTGCTGATGATCGTGAATTTCTTCAAGGC 3207
Qy 3061 CTGCTGAGGCGCGGAGACTTTGATCCGCTCGCAGATGTCGCAACCCCACTGCTGTT 3120
Db 3208 CTGCTGAGGCGCGGAGACTTTGATCCGCTCGCAGATGTCGCAACCCCACTGCTGTT 3267
Qy 3121 CGCTGGATGGATCTCTGAGCCAGACGATAGGGTATGCCAATGTTGTGGCCAAAGTC 3180
Db 3268 CGCTGGATGGATCTCTGAGCCAGACGATAGGGTATGCCAATGTTGTGGCCAAAGTC 3327
Qy 3181 AACGGCCAGATCCGCCCAATCGTGTGGTGACCGCTCCGTTGAGTCTGTCCACGCAACC 3240
Db 3328 AACGGCCAGATCCGCCCAATCGTGTGGTGACCGCTCCGTTGAGTCTGTCCACGCAACC 3387
Qy 3241 GCAGAAAAGGAGATTTCTCCAAAGGGCCATGTTGCTGCACCATTCGCTGGTGTGTC 3300
Db 3388 GCAGAAAAGGAGATTTCTCCAAAGGGCCATGTTGCTGCACCATTCGCTGGTGTGTC 3447
Qy 3301 ACTGTGACTGTTGCTGAGGTGATGAGTCAAGGCTGGAGTGCAGTGCATCATCGAG 3360
Db 3448 ACCGTGACTGTTGCTGAGGTGATGAGTCAAGGCTGGAGTGCAGTGCATCATCGAG 3507
Qy 3361 GCTATGAGTGAAGCAACAACTCACTGCTCTGTTGACGGCAAGATTTGAACGGTGTG 3420
Db 3508 GCTATGAGTGAAGCAACAACTCACTGCTCTGTTGACGGCAAAATCGATCGGTTGTG 3567
Qy 3421 GTTCCTGCTCAACGAAGGTGAAGTGGCGACATTGATGTCGTCGTTTCCTAA 3474
Db 3568 GTTCCTGCTCAACGAAGGTGAAGTGGCGACATTGATGTCGTCGTTTCCTAA 3621

RESULT 3
US-09-134-001C-591
; Sequence 591, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GPC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 591
; LENGTH: 3465
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-591

Query Match 13.7% Score 476.6; DB 4; Length 3465;
Best Local Similarity 48.2%; Pred. No. 2e-122;
Matches 1655; Conservative 0; Mismatches 1724; Indels 54; Gaps 9;
Qy 84 ATTCAAAAGATCTTGGTAGCAACCGCGGCAAAATCGCGTCCGCTGTTTCCGTCGAGC 143
Db 24 AATAAGAAATTAATCTGTTGCTACCGTGGTGAATCGCCATTAGAAATTTAGAGCGGC 83
Qy 144 ACTCGAAACCGGTGCGACCGAGTAGCTATTTACCCCGTGAAGATCGGGGATCATTTCCA 203
Db 84 AGCAGAAATTAATATCATGATACAGTAGCAATTTATTTCTAAATGAAGATAAAAGTTCGTTACA 143
Qy 204 CCGCTCTTTGCTTCTGAAGTGTCCGATTTGCTGAGGCTCACCAGTCAAGGCTA 263
Db 144 TAGATATAAAGCAGATGAATCCTATCTAGTTGGAAGTATTTAGGACCTGCTGAAAGTTA 203

Qy 264 CTTGGACATCGATGAATAATTATCGGTGCGAGCTAAAAAGTTAAAGCAGATGCTATTACC 323
Db 204 TTTGAATATCGAACGATATCATCGAAGTAGCTCTTCGGCAGGTGTCGATGCAATTATCC 263
Qy 324 GGGATATGGCTTCTGCTCAAAATGCCAGCTTGCCTGGAGTGCCTGGGGAACGACAT 383
Db 264 TGGGTATGGTTTTTTAAGTGAAAATGAACAAATTTGACGCGGATGTCTGAGGAAGCAT 323
Qy 384 TACTTTTATTTGGGCCAACCCAGAGGTTCTTGATCTCACCGGTGATAGTCTCTCGTGGGT 443
Db 324 TAAATTTATAGTCCGATCTTGAACATCTAGACATGTTTGGAGATAGGTTAAGGCTAG 383
Qy 444 AACCGCGCGAAGAGGCTGCTGCGAGTTTGGCGG---AATCCACCCCGACAAAAA 500
Db 384 AACAACTGCTATTAAACGCTAACTTACCTGTAATCCCGGTTACAGATGGTCTCTATTGAAG 443
Qy 501 CATCGATGACATCGTTAAAGCGCTGAAGCCAGACATTTACCCCATCTTTGTAAGGCAGT 560
Db 444 TTTTGAAGCTGCGAGACAGTTTCTATGAAGCAGGTTACCCACTTATGATTAAGGCCAC 503
Qy 561 TGGCGGTGGTGGCGGACCGGCTATGCGCTTTGTTCTTCACTGATGAGCTCCGCAAAAT 620
Db 504 AAGCGGTGGCGGTGATAAGGATGCGAATCGTTCGTGAATCAAGCGAATTAGAAGCGC 563
Qy 621 GGCAACAGAAAGCATCTCGTGAAGCTGAAGCGCATTCGGCGAGCGTTTCGGTATATGCGA 680
Db 564 TTTCCATCGTCGCAAAATCAAGACGCAAAAGTCATTTGGTAAATAGCGAAGTTTATATCGA 623
Qy 681 ACGTGCTGTGATTAACCCCGCAGCACATTTGAAGTGCAGATCTTTGGCGATCGCACTGAGA 740
Db 624 AAGATATATTGATAATCCAAAGCATATAGAGGTTCAAGTTATTGGTGATGAATTCGGGAA 683
Qy 741 AGTTGTACACCTTTATGAACGCTGACTGCTCACTGCAGCGCTCGTCACCAAAAAGTTGCGA 800
Db 684 TATCATTTCTTGTGAAGAGATTGCTCCGTACAAGACGTCATCAAAAGTTGTTGA 743
Qy 801 AATTGCCAGCAGCAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAAT 860
Db 744 AGTTGCACCTTTCAGTAGTCTTTCTAAACAAATTAAGAGAGCGAATTTGTGATCGCGAAT 803
Qy 861 AAAGTTCTGCGCTCCATTTGTTTACAGGCGCGGGAACCGTGGAAATCTTTGGTCGATGA 920
Db 804 TCAACTGATGAAAAATATAAAAAACGTCACAGCTGGAACAGTAGAATTTTGTAGTTCTG- 862
Qy 921 AAAGGCAACACCGTTTTCATCGAAATGAACCCACGCTATCCAGTTGAGCACACCGTGAC 980
Db 863 --GGGATGAATTTTCTTCAATTGAGGTTAATCCAGCTGTTCAAGTTGAGCATACAATTAC 920
Qy 981 TGAAGAAGTCAACCGAGGTGGACCTGGTGAAGGCGCAGATCGCTTGGCTGCTGTCGCAAC 1040
Db 921 TGAATGATTACTGTTGATAGACATTTGTGAAAACGCAAAATTTAGTTGCTAATGGAGAATC 980
Qy 1041 CTTGAAGGAATTTGGTCTGACC-----CAAGATAGATCAAGACCCACGCTGAGC 1091
Db 981 GTTATTTGGAGATAAAATCTTATGCCACAGCAAAATGAAAATTCAAACATTAGGGTATGC 1040
Qy 1092 ACTGCACTGCGCATCACACCGAAGATCCAAACACCGCTTCGCGCCAGATACCGGAAC 1151
Db 1041 GATACATGCTGTAACAACCTGAAGATCTCTAATGATTTTATGCCAGATTTGCGCAC 1100
Qy 1152 TATCACCGGTACCGTTCACAGCGGAGCTGGCGTTTCGTTTCGTGACGGTGCAGTCACT 1211
Db 1101 AATTATTGCATATCGATCAAGTGGCGTTTTTGTGTGAGACTTGTATGTCAGGGGATGATT 1160
Qy 1212 C---GGTGGGAAATCACCGCACACTTTGACTCGATGCTGCTGTAAGATGACCTGCCGTGG 1268
Db 1161 CCAAGGTGCAAAATTTTCACTTACTACGATTCATCTATTAGTTAAGCTTTTCTACACATGC 1220
Qy 1269 TTCCGACTTTTGAACCTGCTGCTCGTGACACGCGCGTTCGCTGAGTTTCAACGCTGC 1328
Db 1221 CGTTTCATTTTAAACAAGCTGAAGAAAAATGGAAACGTTTATTACCGCAAAATGCAATTCG 1280

QY 1329 TGGTGTGCAACCAATTTGGTTCTTCTGCGTGGCTTCTGCGGAGAGGACTTCACTTC 1388
DB 1381 TGGCGTAAAGACGAATATTCATTTCTCATCAATGTTATGCGTAATGATATAATTTAGAAG 1340
QY 1389 CAAGCGCATCCGACCGGATTTATCGCGATACCCACACTCTCTTCAGGCTCCACTGC 1448
DB 1341 TGGTGATTTATCTACTATAATTTATTTGAAGAAACACCTGGAATCTTTGATATTCGACCGAC 1400
QY 1449 GGATGATGAGCAGGAGCGCATCTGGAATTTAGTGGAGATGTTGCGAGATGTTCAACGTTGAACAGCTCA 1508
DB 1401 ATTGGACAGAGGTACCAAGACTTTAGAGTATATTTAGTAAATGTTGACGATAAACGGAATTTCC 1460
QY 1509 TGGTGT-----GGTCCAAAGATGTTGAGCAGCAATCGATAAGCTGCGCAACATC-- 1560
DB 1461 TAATGTAGAAAGCGTCCAAACACAGAAATATGAATCTACCAAAATCCCAAAATTTCTCA 1520
QY 1561 -AAGGATCGCACTGCCACCGGTTCCGTTGACCGCTGAAGCAGCTTGGCCGAGCGC 1619
DB 1521 AARGAAATCAATCAGTTATTTGGAAACAAACAAATTTCTTGAGCAACATGGACCAACAGG 1580
QY 1620 GTTTGCTCGTATCTCGTGAGAGGAGCGACTGGCAGTTACTGATACCACTTCCGCA 1679
DB 1581 TGTACAAATTTGGTTAGAGAACAGAAAGATGTTTAAATACCGTACTACATTTAGAG 1640
QY 1680 TGCACACAGTCTTGTGTGGACCCGAGTCCGCTCAATCGCACTGAAGCTGCGCGAGA 1739
DB 1641 TGCACACCAATCTTTACTTGAACACGTTGAAGAACAAAGATATGATGAACATTTGCATC 1700
QY 1740 GGCCGTGCAAGCTGACTCTCTGAGCTTTTCTCGTGGAGCGCTTGGGCGCGCGACCTA 1799
DB 1701 TAAACTCTGTAAGTTTAAAGATAGTTTTCATTAGAAATGTTGGGTGGTGCAACAT 1760
QY 1800 CGATGTGGCGATCGCTTCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTTGC 1859
DB 1761 TGATGTGCTATAATTTCTTGAAGAGAAATCCATGGGAACGTTTAGAAAGATTCGCA 1820
QY 1860 GCGGATGCGGAATGAACATTTCAATGCTGCTTCCGCGCGCAACACCGTGGGATACAC 1919
DB 1821 AGCCATTCGAATGTTTATCCAAATGTTTACGAGCTTCGAAGCGAGTAGGTTATAA 1880
QY 1920 CCGGTACCGAGCTCGCTCGCGCTGTTGTTAAGCAAGCTGCACTCGCGCTGGA 1979
DB 1881 AAATATCTGATATGTTAAATGAAGAAATCGTTCATGAAGTGCAGAAAGCTGGTGA 1940
QY 1980 CATCTTCGATCTTCGACGCGCTTAAAGAGCTTCCAGATGCTCCAGCAATCGAGC 2039
DB 1941 TGTTCGCTATTCGACTCAATGAACTGGGTTGATCAATGAAGTAGCGATGAAGC 2000
QY 2040 AGTCTGGAGACAACACCGCGGTAGCGAGTGCTATGCGTTATCTGTTGATCTCTC 2099
DB 2001 TGTTCAGAAAGCTGGAATGTTATCTGAGGGTACAATTTGCTATACAGGTGATTTTAA 2060
QY 2100 TGATCCAAATGAAAGCTCTACACCTTGGATTAACCTTAAGATGGCAGAGGAGATCGT 2159
DB 2061 TGTGAAACGTTCCATATTTATCTTTAGATTTATACGTTAAATGGCTAAAGAACTGGA 2120
QY 2160 CAAGTCTGGCGCTCACATCTGGCCATTAAGGATATGGCTGCTTCTGCGCCAGCTGC 2219
DB 2121 AAGAGAAGGATCCATATATTAGCAATTAAGATATGGCTGTTTATTGAAACCGAAAGC 2180
QY 2220 GGTAAACAGCTGTCACCGCACTGCGCGCTGAATTCGATCTGCCAGTGCAGTGCACAC 2279
DB 2181 AGCTTACGAATTAATTTGGTGAATTTAGTGGGCAACACATCTTCCAATTCATTACATAC 2240
QY 2280 CGAGCACTGCGGTGGCGGCTGCTACTTCTGCTGAGCTCAAGCTGGTGGAGA 2339
DB 2241 ACATGATCTAGTGAATGATTTGACATATAAACAAGCAATTTGATGCTGGCGTGA 2300
QY 2340 TGTCTGTGACGTTGCTCGCACCACTGCTGCGACCACTCCAGCCACTCCCTGCTGCG 2399
DB 2301 TATTATAGATCTGCTGTTGCTGTTATGAGTGGTTTAAAGTCAACCAAGTGCAATTC 2360
QY 2400 CATGTTGTGCAATCGCGCAACCCGTCGCGATACCGGTTTGGAGCTCGAGCTGTTTC 2459

DB 2361 ATTATATTATGCACTAAATGGATTTCCACGTAATTTAAGAACTGATATTTGATGGTTAGA 2420
QY 2460 TGACCTCGAGCGGTACTGGGAAGCTGTGCGGAGACTGTACCTGCCATTTGAGTCTGGAAC 2519
DB 2421 AGAGTTGAGTCAATTACTTGGTCTGTAGTCAGACCTTACTATGCGAGCTTTGAGAGTGAT 2480
QY 2520 CCCAGGCCCAACCGGTGCGCTTACCGCCAGGAATCCAGGCGGACAGTTGTCCAACCT 2579
DB 2481 CAATCAACAAATACAGNAATTTATCAACATGAATGCCAGTGGCCCAATATTCAAACT 2540
QY 2580 GCGTGCACAGGCAACCGCACTGGGCTTGTGATCGTTCAGCTCATCGAAGCAACTA 2639
DB 2541 AAGTCAACAAAGTAAAGTTTAGGATTTGGCGAAGCTTTTGTGAAGTCAAGAGATGA 2600
QY 2640 GCGAGCCGTTAATGAGATGCTGGGAGCGCCAAACCAAGGTCAACCTCTCAAGTTGT 2699
DB 2601 TCGTGTGTCACATCTCTCTTTGGAGATCTTTGAAAAGTAACCACTTCCAAGGTAGT 2660
QY 2700 TGGCGACCTCGCACTCCACCTGGTTGGTGGGCTGTAGATCCAGCAGACTTTTGTGCGAGA 2759
DB 2661 TGGAGATATGGCACTATATATGTTGCAAAATGATCTTGTGAAGATACGGTCACTAATGA 2720
QY 2760 CCCACAAAGTACGACATCCAGACTCTGTCAATCGCGTTCTTGGCGCGGAGCTTTGGTAA 2819
DB 2721 TGGTTATAAATTTAGATTTCCCGAATCTGTTGTCTATCTTTAAAGGTGACATTTGGACA 2780
QY 2820 CCTCCAGGTGCTGGCCAGAACCACTCGCACCCGCACTGGAAAGCGCGCTCCGAAGG 2879
DB 2781 ACTGTCAACGATTTCAACAGAAATTCGAAGATGTTATTTAAAGGACAGCAACCAAT 2840
QY 2880 CAAGGCACTCTGACGGAGTTCTTGAGNAGAGCAGCGCACTCGACGCTGATGATTC 2939
DB 2841 TACTGAAGACAGGTGAATCTTGGAGCGGTCAATTTTGAAGCAATCCGTCAAGAAT 2900
QY 2940 CAAGA-----AGTCCGAACAGCTCAACCGCTCTGCTCTCC 2978
DB 2901 AAGCGCATACAAACAGAGAGGTAAACAGAAAGATATAATTTAGTTATGTTACTTATCC 2960
QY 2979 GAAGCAACCGAAGAGTTCTCTGAGCACCGTCCGCTTTGGCAACACCTCTGCGCTGA 3038
DB 2961 GAAGGTATATAACAAATATATTTCAACAGAAAGCAATTTGGTAATGTTATCTTTACTGA 3020
QY 3039 TGATCTGAATTTCTTCTACGAGTGTGTCAGGCGCGCGAGACTTTGATCCGCTGCGCAGA 3098
DB 3021 TACACCGACATTTCTTATTTGGCATGCTAATG---GTGAAACAGTTGAAATTTGA 3077
QY 3099 TGTGCGCACCCCACTGTTGTTGCTGATGCGATCTCTGAGCCAGACGATAAGGTTAT 3158
DB 3078 TACTGTTAAAGCTTAATTTAATTTAGAAACAATCAGTGAACCAAGATGAAATGTTAA 3137
QY 3159 GCGCAATGTTGGCCCAACGTCAGCGCCAGATCCGCCCAATCGGTGTCGTCGACGCTC 3218
DB 3138 GCTACAAATTTATACGCTATGATGTTCAAGCAAGACGTTATTTCAAGATGAAAA 3197
QY 3219 CGTTGAGTCTGTCAACGCAACCGCAAGGAGGAGGATTCCTCAACAGGCGCATGTTGC 3278
DB 3198 TGTAAACCAATGCTTAATGTTAACTAAGCGGTAATCAATCCAAATCATATGG 3257
QY 3279 TGACCAATTCGCTGGTGTGTTCTACT---GTGACTGTTGTGAGGTGATGAGTCAAGC 3335
DB 3258 TGTCAAAATGCTGTTGTTGTAATCAAGTCAAGTGTCTGTAGGCGATGAGTTCAAGC 3317
QY 3336 TGGAGATGCACTGCAATCATCGAGGCTATGAAGATGGAAGCAACATCACTGCTCTGT 3395
DB 3318 TAATCAGCCATTTAATCACTGAAGCAATGAAGATGGAACACGATTCAGGCACAT 3377
QY 3396 TGACGCAAGATTTGAACCGCTTGTGTTCTGCTGCAACGAAGGTGGAAGGTGCGACTT 3455
DB 3378 TGTGGAATTTATTAACAAATCAATGTTGTTAATGAGATGCCATTTGCCACAGGATTT 3437
QY 3456 GATCGTCTGCTT 3468

Db 3438 ATTAGTGGAAATT 3450

RESULT 4

US-07-956-700B-5

; Sequence 5, Application US/07956700B

; Patent No. 5539092

; GENERAL INFORMATION:

; APPLICANT: Robert Haselkorn and Piotr Gornicki

; TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA

; TITLE OF INVENTION: Carboxylase

; NUMBER OF SEQUENCES: 116

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Arnold, White & Durkee

; STREET: 321 No. 5539092th Clark Street

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60610

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy Disk

; COMPUTER: IBM PC Compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII-DOS

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/956,700B

; FILING DATE: 19921002

; CLASSIFICATION: 800

; ATTORNEY/AGENT INFORMATION:

; NAME: Thomas E. No. 5539092thrup

; REGISTRATION NUMBER: 33,268

; REFERENCE/DOCKET NUMBER: ARCD:058

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 1-312-744-0090

; TELEFAX: 1-312-755-4489

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1362 base pairs

; TYPE: Nucleic acid

; STRANDEDNESS: Single

; TOPOLOGY: Linear

; MOLECULE TYPE: Oligonucleotide

; US-07-956-700B-5

Query Match

Best Local Similarity 53.6%; Pred. No. 2.3e-75;

Matches 714; Conservative 0; Mismatches 604; Indels 15; Gaps 3;

Qy 85 TTCAAAAGATCTTGGTAGCAAAACCGGGCGAAATCGCGTCTTTCGGTGCAGCA 144

Db 7 TTCAACAGATCTTGATCGCCAATCGCGCGAAATCGCCCTCGCGACTTGT 66

Qy 145 CTCGAAACCGGTGCAGCCACCGTAGCTATTATACCCCGTGAAGATCGGGGATCTCCAC 204

Db 67 CAAGACTCGGGATCGGCACGATCGCGTTCACTCACTGTGGATCGCAACGCGCTCAT 126

Qy 205 CGCTCTTTTGGTCTCTGAAGTGTCCGCAATGGTACTGAAGGCTCACCAAGTCAAGGCGTAC 264

Db 127 GTGCAGTTAGCGGACGAAGCGTCTGTATTGG---CGAAGCGCCAGCAGCAAAAGCTAT 183

Qy 265 CTGACATCGATGAATATATCGGTGCAGCTAAAGAAAGTTAAAGCAGATGCTATTACCG 324

Db 184 CTCATATCCCAACATCATTTGCGGCGCCCTGACCCCTTAATGCGAGCGCCATTCACCCC 243

Qy 325 GGATATGCTTCTGCTGTAAGAAATGCCAGCTTCCCGCGAGTCCGGGAAACGGCAT 384

Db 244 GGCTATGCTTCTTGGCGGAGAAATGCCCTTTGAGAAATCTGCGCGGATCAATCTC 303

Qy 385 ACTTTTATTTGCCCAACCCCAAGAGTTCTTGATCTCACCGGTGATAAGTCTCGTGC---G 441

Db 304 ACCTTTATTTGCCCGCCCGCCGATTCGATTCGAGCCATGGCGGATNAATCACCGCTAAG 363

Qy 442 GTAAACCGCCGGAAGAGGCTGGTCTGCCAGTTTGGCGGAATCAACCCCGCAAAAC 501

Db 364 GAAACAATGACGCGGTCCGGCTTCGACGATTCGGGCGAGTGACGGTCTCTCAGCGAT 423

Qy 502 ATCGATGACATCGTTAAAAAGCGCTGAAGCGCAGACTTACCCCATCTTTTGTAAAGCGAGTT 561

Db 424 GTTGATTCGGGTGCCAAAGTTGCTCGCGAGATCGGCTATCCCGTCATGATCAAGCGACG 483

Qy 562 GCCGGTGTGGCGAGACGCGGTATGCGCTTTTGTTCCTCACTGATGAGCTCCGCAAAATTG 621

Db 484 GCGGGGCGGTGCTGCGGTATGCGGCTGCTGCGTGCACCTCGCAGATCTGGAAGAACTG 543

Qy 622 GCAACAGAAGCATCTCGTGAAGCTGAAGCGCATTCGCGAGCGGTTCGGTATATGTGAA 681

Db 544 TTCCTTCTGCCCAAGGAGAGCCGAGCAGCTTTTGGGAATCCAGGACTGTATCTCGAA 603

Qy 682 CGTCTGTGATTTAACCCCGCAGCAATTTGAAGTGCAGATCTTGGCGGATCGCACTGGAGAA 741

Db 604 AATTTATCGATCGCCACCGCCAGTTTGAATTTTCAGATCTTGGCGGATCGCTACGGCAAT 663

Qy 742 GTTGTAACCTTTATGAACGTGCTCTCACTGAGCGGTGCTACCAAAAAAGTTGTGAA 801

Db 664 GTAGTGCATCTAGCGGAGCGGATTTGCTCCATTCAGCTGCTCAACAAAAGCTGCTCGAA 723

Qy 802 ATTGCGCCAGCAGCATTTGGATCCAGAACTGCGTGCATCGCATTTGTGCGGATGCACTA 861

Db 724 GAAGCCCCAGTCCGCGCTATCGGCAGACCTCGCGCAGAAATGGCGATGCGCGCGTC 783

Qy 862 AAGTTCTGCGCGCTCCATTGGTTTACAGGGCGCGGAAACCGTGGAAATTTCTTGGTCCATGAA 921

Db 784 AAGTCTGCTCAAGCGATCGGTACATCGGTGCGGCGACCGTGGAGTTTCTTGGTGCATGCG 843

Qy 922 AAGGGCAACCAAGTTTTCATCGAAATGAACCCAGTATCCAGGTTGAGCACACCGTGAAT 981

Db 844 ACCGCAACTTCTACTCATGAGATGAATACCGCATCCAAAGTCGAGCATCCAGTCAACA 903

Qy 982 GAAGAGTCAACGAGGTGGACCTGGTGAAGCGCAGATGGCTTGGTGTGTTGGTGCACAC 1041

Db 904 GAATGATTAAGGACTGGACTTGAATTCGGGAGCAGATTCGGATTGGCCCAAGGCG- 958

Qy 1042 TTGAAGGAATTTGGTCTGTGACCCCAAGATAAGATCAAGACCCACCGTGCGACACTGCAGTGC 1101

Db 959 ----AAGCGTTCGCTTCGCGCAAGCCGATATTCACCTGCGGCGCCATGCGATCGAATGC 1014

Qy 1102 CGCATACACCGAAGATCAAAACACCGGTTCGCGCCAGATACCGGAATATCACCGCG 1161

Db 1015 CGTATCAATCGGAAGATCGGAATACAAATTCGCGCGGAATCTCGCGCGCATTTACAGGC 1074

Qy 1162 TACCGCTCACGAGCGGAGCTGGGTTTCTGTTGACCGGTGCGAGCTCAGCTCGGTGGCGAA 1221

Db 1075 TATTTACCGCGCGCGCGCGGTTCTGTTGCGATTCCCATGTTTATACCGACTACGAA 1134

Qy 1222 ATCACCGCACACTTTGACTCCATGCTGTTGAATGACCTGCGCGGTTCGAGCTTTGAA 1281

Db 1135 ATTCCGCGCTTATACGATTCGCTGATTTGGCAATGATTGTTGCGGTGCAACAGGGAA 1194

Qy 1282 ACTGCTGTGCTCGTGCAAGCGCGGTGCTGAGTTCAACCGTGTCTGGTGTGTTGCAACC 1341

Db 1195 GAGCGCATCGCGGATGCGAGGTGCTCTCGGGAATGCGCCATCACCGGCTTCCCGACG 1254

Qy 1342 AACATTTGTTTCTTGGCTGCTGCTCGGGAAGAGACTTCACTTCCAGCGCATGCC 1401

Db 1255 ACCCTTAGTTTCCATCAGCTGATGTTGTCAGATGCTGAGTTTCTGCGCGGGAACTCTAT 1314

Qy 1402 ACCGATTTATCG 1414

Db 1315 ACCAACTTTGTTG 1327

RESULT 5

US-08-476-537-5

; Sequence 5, Application US/08476537

; Patent No. 5756290

; GENERAL INFORMATION:

APPLICANT: Robert Haselkorn and Piotr Gornicki
 TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
 NUMBER OF SEQUENCES: 116
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Arnold, White & Durkee
 STREET: 321 No. 5756290th Clark Street
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60610
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy Disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII-DOS
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/476,537
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/956,700
 FILING DATE: 10/21/92
 ATTORNEY/AGENT INFORMATION:
 NAME: Thomas E. No. 5756290thrup
 REGISTRATION NUMBER: 33,268
 REFERENCE/DOCKET NUMBER: ARCD:058
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 1-312-744-0090
 TELEFAX: 1-312-755-4489
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1362 base pairs
 TYPE: Nucleic acid
 STRANDEDNESS: Single
 TOPOLOGY: Linear
 MOLECULE TYPE: Oligonucleotide

Query Match 8.8%; Score 306.6; DB 1; Length 1362;
 Best Local Similarity 53.6%; Pred No. 2.3e-75;
 Matches 714; Conservative 0; Mismatches 604; Indels 15; Gaps 3;

85 TTCAAAAGATCTTGGTAGCAACCGCGGGAATCGCGTCCGCTTCCGTCGACGA 144
 7 TTCAACAGATCTGATCGCCATCGCGGGAATCGCGCTCGCATCTCGCATTTGT 66
 145 CTCGAACCGGTGCGACCGTAGCTATTACCCCGTGAAGATCGGGGATTCACAC 204
 67 CAAGAACTCGGATCGGACGATCGCGCTCACTCACTGTGGATCGCAACGCGTCC 126
 205 CGCTCTTTTGGCTTCTGAAGCTGTCGCAATTGGTATGAAGGCTCACAGTCAAGCGGTAC 264
 127 GTGACGTTAGCGGACGAACGCGTCTGTATTGG---CGAAGCGGCGAGCGCAAAAGCTAT 183
 265 CTGGACATCGATGAATATTCGGTGCAGCTAAAGATTTAAAGCAGATGCTATTACCGG 324
 184 CTCAATATCCCAACATCATTTGCGGCGGCGCTGACCCCTTAATGCGGCGGCAATTCACCCC 243
 325 GGATATGGCTTCTGCTGAATATCCAGCTTGCCTGCGGAGTGGCGGAGTGGCGGAAACCGCAT 384
 244 GGCTATGGCTTCTTGGCGGAGATGCCCTTTGCAAGAAATCTGCGCGCATCACCATCTC 303
 385 ACTTTTATTTGGCCCAACCGCAGAGTCTTGTATCTCACCGGTGAATGCTGTC---G 441
 304 ACTTTTATTTGGCCCAACCGCAGAGTCTTGTATCTCACCGGTGAATGCTGTC---G 363
 442 GTAACCGCGGGAAGAGCTGCTGCGAGTTTGGCGGAATCCACCCCGAGCAAAAC 501
 364 GAACAAATGACGCGGTGCGCTGCGAGTTCCGAGATTCGCGGCGAGTACGCTGCTGACGAT 423
 502 ATCGATGACATCGTTAAAGCGCTGAAGCCAGACTTACCCATCTTTGTAAGGCGATT 561

424 GTTGATTGGCTGCCAAAAGTTGCTGCCGAGATCGGCTATCCCGTCATGATCAAAAGCGACG 483
 562 GCCGTGTGTGCGGACGCGGTATCGGCTTTGTTCTTCACTGATGAGCTCCGCAAAATTG 621
 484 CGGGGGCGGTGTGCGGTATCGGCTGTGCGTGCAGATCTGGAAGAACTG 543
 622 GCAACAGAAAGCATCTGTAAGCTGAAGCGCATTCGCGGACGTTTCGGTATATGTCGAA 681
 544 TTCCTGTGCTGCCAAGGAGAGCGAGCGAGCTTTTGGGAATCCAGGACTGTATCTCAA 603
 682 CGTGTGTGATTAACCCCGACGACATTTGAAGTGCAGATCTTGGCGATCGCACTGGGAA 741
 604 AAATTTATCGATCGCCACGCGCATTTGAATTCAGATCTTTGGCGATCGCTACGCGCAT 663
 742 GTTGTACACCTTTATGAACGTGACTGCTCACTGAGGCTGCTCAACAAAAGATTGTGAA 801
 664 GTAGTGCATCTAGGCGAGCGGATTTGCTCCATCAACGCTGTCACAAAAGCTGCTCGAA 723
 802 ATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGTATCGCATTTTGGCGGATCGAGTA 861
 724 GAAGCCCCAGTCCGCGCTATCGGACAGCTCGCGCAGAAAATGGGCGATGCCGCCGTC 783
 862 AAGTTCTGCGCGCTCCATTGGTTTACCGGGCGCGGAAACCGTGGAAATTTCTTGTGATGAA 921
 784 AAGTGTCTCAAGCGATCGGCTACATCGGTGCGGCAACCGTGGAGTTTCTGTCGATGCG 843
 922 AAGGCAACACCGTTTTCATCGAAATGAACCCAGCTATCCAGGTGAGCACACCGTGAAT 981
 844 ACCGCAACTTCTACTTTCATGAGATGAATACCCGATCCCAAGTCGAGCATCCAGTACA 903
 982 GAAGAAGTCAACGAGGTGAGCTGTGAAGGCGGAGATCGGCTGCTGCTGCTGCTGCTGCTG 1041
 904 GAAATGATTACGGGACTGGACTTTGATTTGCGGAGCAGATTCGGAATGCCAAGCGG 958
 1042 TTGAAGAAATTTGGTCTGACCCAGATTAAGATCAAGACCCAGGTGAGCACACCGTGAAT 1101
 959 ----AAGCGCTGCGCTTCCGCAAGCGGATTTCACTGCGGCGCATGCGATCGAATGC 1014
 1102 CGCATCACCAACCGAAGATCCAAACAAACCGCTTCCGCGCAGATACCGAACTATCACCGG 1161
 1015 CGTATCAATGCGGAAGATCCGAATACAAATTTCCGCGCGAAATCTCGCGCGCATACAGC 1074
 1162 TACCGCTACACGAGCGGAGCTGCGCTTCTGTTGAGGTCGAGCTCAGCTCGGTGGCGAA 1221
 1075 TATTTACCGCGCGGCGCGCTGCTGTCGATTTCCATTTCCATTTTATACCGACTACGAA 1134
 1222 ATCACCGCACACTTTGACTTCCATGCTGTGTTGAGAAATGACCTGCGGTGGTTCCGACTTTGAA 1281
 1135 ATTCCGCCCTATTACGATTTCGCTGATTGGCAAAATGATTGCTGCGGCTGCAACCGGAA 1194
 1282 ACTGCTGTTGCTGTCGACAGCGCGGTTGGCTGAGTTCACCGTGTCTGCTGTTGCAACC 1341
 1195 GAGCGCATCGCGGATGACGCTGCTTCCGGAATGCGCCATCACCGGCTTGGCGAGC 1254
 1342 AACTTGGTTTCTGCTGCTGCTGCGGAGAGGACTTCACTTCCAAAGCGCATCGCG 1401
 1255 ACCCTAGTTTTCATCAGCTGATGTTGAGATGCTGAGTCTGCTGCGGGGAACTCTAT 1314
 1402 ACCGATTTATCG 1414
 1315 ACCAATTTGTTG 1327

RESULT 6
 US-08-485-607-5
 ; Sequence 5, Application US/08485607
 ; Patent No. 5792627
 ; GENERAL INFORMATION:
 ; APPLICANT: Robert Haselkorn and Piotr Gornicki
 ; TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
 ; TITLE OF INVENTION: Carboxylase
 ; NUMBER OF SEQUENCES: 116
 ; CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5792627th Clark Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,607
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/956,700
FILING DATE: 10/21/92
ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. No. 5792627thrup
REGISTRATION NUMBER: 33,268
REFERENCE/DOCKET NUMBER: ARCD:058
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-312-744-0090
TELEFAX: 1-312-755-4489
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1362 base pairs
TYPE: Nucleic acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Oligonucleotide
US-08-485-607-5

Query Match 8.8%; Score 306.6; DB 1; Length 1362;

Best Local Similarity 53.68; Pred No. 2.3e-75;
Matches 714; Conservative 0; Mismatches 604; Indels 15; Gaps 3;

Qy 85 TTCACAAAGATCTTTGGTAGCAAAACCGCGGAAATCGCGTTCGCTTTCCGTCGACGA 144
Db 7 TTCACAAAGATCTTCGCAATCGCGGAAATCGCGTTCGCTTTCCGTCGACGTGT 66
Qy 145 CTCGACATCGATGAAATTTATCGTGCAGCTAAAGGTTAAAGCAGATGCTATTTCACCG 324
Db 184 CTCGATATCCCAACATCATTTTCGCGGCGCTGACCCCTATATGCGGCGCATTCACCCC 243
Qy 325 GGATATGCTTCTGTTGAAATATCGCGGCGAGTGGCGGAGTGGCGGAAACCGGATT 384
Db 244 GGCTATGCTTCTTTGGGAGAAATCGCGCTTTGCAGAAATCTGCGCGCATCATCTC 303
Qy 385 ACTTTTATGSCCAACCCAGAGTTCTTGATCTCACCGGTGATAGTCTGTCG---G 441
Db 304 ACTTTATGSCCCAGCCCGATTCGATTCGAGCCATGGCGGATTAATCACCAGTAAG 363
Qy 442 GTAAACCGCGGAGAGGCTGGTCTGCCAGTTTGGCGGAATCCACCCCGAGCAAAAC 501
Db 364 GAAACAATGACGGGTTCGGGTTCCGACGATTCGGGCGATGACGCTGCTGACGGAT 423
Qy 502 ATCGATGACATCGTTAAAGCGCTGAAGCGCAGACTTACCCCATCTTTTGTAAAGCGAGTT 561
Db 424 GTTGATTCGGTCCAAAGTTGCTGCGGAGATCGGCTATCCCGTCATGATCAAGCGACG 483
Qy 562 GCGGCTGGTGGCAGCGGATGATCGGCTTTGTTCTTCACTGATGAGCTCCGCAATG 621
Db 484 GCGGGGGCGGTGTCGCGGTATGCGGCTGGTGTGACCTCGAGATCTGGAAAACTG 543

Qy 622 GCAACAGAGCACTCTGTGAAGCTGAAGCGGCATTCGGCGACGGTTCCGTATATGTGAA 681
Db 544 TTCCTTGTCTGCCAAGGAGAACGCCAGCGACCTTTTGGGAATCCAGGACTGTATCTCGAA 603
Qy 682 CGTCTGTGATTAAACCCCGAGCATTGAAGTCAGATCTTGGCGATCGCACTGAGAGAA 741
Db 604 AATTTATCGATCGCCACCGCACGTTTGAATTTTCAGATCTTGGCCGATGCCCTCGGCAAT 663
Qy 742 GTTGACACCTTTATGAACGTGACTGCTCACTGAGCGGTGCTCAACCAAAAAGTTGTGAA 801
Db 664 GTAGTGCATCTAGCGAGCGGATTTCTCCATTCAAGTCTGTCACCAAAAAGTCTCTCGAA 723
Qy 802 ATTGCGCAGCAGCAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGGGATGCGAGTA 861
Db 724 GAAGCCCGAGTCGCGGCTATCGGCAGACCTGCGGAGAAATGGCGCATGCGCGCGTC 783
Qy 862 AAGTTCTGCGCTCCATTGTTTACCGGCGCGGAGACCGTGGAAATCTTGGTTCGATGAA 921
Db 784 AAGTCTGCTCAAGCGATCGGCTACATCGGTGCGCGCACCGTGGAGTTTCTGGTTCGATGCG 843
Qy 922 AAGGCAACCAACGCTTTTCATCGAAATGAACCCAGTATCCAGGTTGAGCACACCGTGAAT 981
Db 844 ACCGCAACTTCTACTTCATGGAGATGAATACCCGCATCCAAGTCGAGCATCCAGTCACA 903
Qy 982 GAAGAAGTCAACGAGGTGACCTGGTGAAGCGCAGATGCGCTTTGGTGTGTTGCAACC 1041
Db 904 GAAATGATTACGGGACTGGACTTGATTGCGGAGCAGATTGGATTGCCCAAGCGG---- 958
Qy 1042 TTGAAGGAATTTGGTCTGACCCCAAGATAAGATCAAGACCCAGTCAGGTCAGTGC 1101
Db 959 ----AAGCGCTGCGCTTCCGCGAAGCGCATATTCACTGCGCGCCATGCGATCGAATGC 1014
Qy 1102 CGCATACACGAGGATCAAAACCAACGCTTCGCGCCAGATACCGGAATCTATCACCGCG 1161
Db 1015 CGTATCAATCGGAGATTCGGAATACAATTTCCGCGCGAATCTGCGCGCATTTACAGGC 1074
Qy 1162 TACCGCTCACAGCGGAGTGGCGTTCTGTTTACCGGTGCGCTCAGCTCGGTGGCGAA 1221
Db 1075 TATTTACCGCGCGCGCGCTGCTGTTGTTGATTCCTCATGTTTATACGACTACGAA 1134
Qy 1222 ATACCGCACACTTTGACTTCATGCTGGTGAAGATGACCTGCGGTGGTTCGGACTTTGAA 1281
Db 1135 ATTCCGCGCTTATACGATTCGCTGATTGGCAAAATGATTGCTGGGGTGCACACCGGAA 1194
Qy 1282 ACTGCTCTGCTGCACAGCGCGTTCGCTGAGTTTCAACCGTCTGCTGTTGCAACC 1341
Db 1195 GAGCGATCGCGGATGCGAGCTGCTCTCGGGAATGCGGCATCACCGCTTCCCGACG 1254
Qy 1342 AACATTGGTTTCTTGGCTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAAGCGCATCGCC 1401
Db 1255 ACCCTTAGTTTCCATCAGCTGATGTTGCGAGATGCTTCTGCGGGGAACTCTAT 1314
Qy 1402 ACCGATTTCG 1414
Db 1315 ACCAACTTTGTTG 1327

RESULT 7

US-08-475-879-5
; Sequence 5, Application US/08475879
; Patent No. 5972644
; Patent No. 5972644 5786170
; GENERAL INFORMATION:
; APPLICANT: Robert Haselkorn and Piotr Gornicki
; TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
; TITLE OF INVENTION: Carboxylase
; NUMBER OF SEQUENCES: 116
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: 321 No. 5972644 5786170th Clark Street
; CITY: Chicago
; STATE: Illinois

COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,879
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/956,700
FILING DATE: 10/21/92
ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. No. 5972644 5786170thrup
REGISTRATION NUMBER: 33,268
REFERENCE/DOCKET NUMBER: ARCD:058
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-312-744-0090
TELEFAX: 1-312-755-4489
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1362 base pairs
TYPE: Nucleic acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Oligonucleotide
US-08-475-879-5

Query Match 8.8%; Score 306.6; DB 2; Length 1362;
Best Local Similarity 53.6%; Pred. No. 2.3e-75;
Matches 714; Conservative 0; Mismatches 604; Indels 15; Gaps 3;

QY 85 TTCAAAAAGATCTTGGTAGCAAAACCGCGGCGAAATCGCGTCCGTGCTTTCGCGTAGCA 144
DB 7 TTCAACAGATCTTGATCGCAATCGCGCGAAATCGCCCTGCGCAATCTCGCACTGT 66
QY 145 CTCGAACCGGTGAGCCACCGTAGCTATTACCCCGTGAAGATCGGGATCATTCAC 204
DB 67 CAAAGATCTGGATCGGACGATCGCGCTTCACTCCACTGTGATCGCAAGCGCTCAT 126
QY 205 CGCTCTTTGCTTCTGAGCTGTGCGATCGGTACTGAGCTCACCAGTCAAGCGCTAC 264
DB 127 GTGCAATAGCGGACAGCGGTCTGTATTTG---CGAAGCGGCGAGCAAGACTAT 183
QY 265 CTGGACATCGATGAATATATGCTGACGCTAAAGATTAAAGCATGCTATTACCG 324
DB 184 TCAATATCCCAACATATTGCGCGGCGCTGACCCCTAATGCGCGCCATTCACCC 243
QY 325 GGATATGGCTTCTCTGTAATAATCCAGCTTGCCGCGAGTGCAGGAAACCGGATT 384
DB 244 GGCTATGGCTTCTTGGCGGAGAAATCGCGCTTTGCGAAGAAATCGCGCGATCACCATCTC 303
QY 385 ACTTTTATGGCCCAACCCAGAGGTTCCTTGATCTCACCGGTGATAGTCTCGTGC---G 441
DB 304 ACTTTTATGGCCCAACCCAGAGGTTCCTTGATCTCACCGGTGATAGTCTCGTGC---G 441
QY 442 GTAACCGCGGCGAAGAGCTGTGCTGCGAGTTTTCGCGGAATCCACCCCGAGCAAAAC 501
DB 364 GAAACAAATCGAGCGGTGCGGTTCGAGGATTCGCGGCGAGTCTGCTGACGAT 423
QY 502 ATCGATGACATGTTAAAGCGCTGAAGCCAGACTTACCCCATCTTTGTAAGCGAGTT 561
DB 424 GTTGATTCGGTGCCTGCTGCTGCGGAGTCTGCGGAGTCTGCGTATCAGTCAAGCGAG 483
QY 562 GCGGTGTGGGCGGAGCGGTATGCGTCTTTCTTCACTGATGAGTCCGCAATTG 621
DB 484 GCGGGGCGGCGGTGCGCGGTATGCGGTGCGGTGAGCTTGCAGATCTGGAANAATG 543
QY 622 GCACAGAAAGCATCTGCTGAAGCTGAGCGGATTCGCGGAGCGTTCGGTATATCTGAA 681
DB 544 TTCTTGTCTGCCCAAGAGGAGCGGCGGCGCTTTTGGGAATCCAGGACTGTATCTGAA 603

QY 682 CGTGTGTGATTAAACCCCGACGACATTTGAAGTGCAGATCTTGGCGGATCGCACTGGAGAA 741
DB 604 AAATTTATCGATCGCCACGCGCACTTTGAATTTTCAGATCTTGGCCGATGCTACGCAAT 663
QY 742 GTTGTACACCTTTATGAACGTGACTGCTCACTGACGCGTGTGACCAAAAGTTGTGAA 801
DB 664 GTAGTGATCTAGGCGAGCGGATTTCTCATTCAGTGTGCTACCAAAAGCTCTCGAA 723
QY 802 ATTGCGGCGACGACATTTGGATTCAGAACTGCTGATCCATTTGTGGGATGCGATGA 861
DB 724 GAAGCCCCAGTCCGGCGCTATCGCGAGACTCGCGGAGAAAAATGGGCGATCCCGCGTC 783
QY 862 AAGTTCTCGCGCTCATTTGGTTTACGAGCGCGCGGAAACCGTGGAAATTTCTTGGTGA 921
DB 784 AAAGTCGCTCAAGCGATCGGCTATCGGTGCGCGCAACCGTGGAGTTTCTTGGTGA 843
QY 922 AAGGCAACCACTTTTCATCGAAATGAACCCAGCTATCCAGGTTTGAGCACACCGTGA 981
DB 844 ACCGCGCAACTTCTACTTTCATGAGATGAATACCCGCAATCCCAAGTCGAGCATCCAGTCA 903
QY 982 GAAGAAGTCAACGAGTGGACTGTGTGAAGCGGAGATGCGCTTGGCTGTCTGGTGCAAC 1041
DB 904 GAATGATTACGAGTGGACTTGAATTCGCGGAGCAGATTCGATTTGCCCAAGCG--- 958
QY 1042 TTGAAGGATTTGGTCTGACCCCAAGATGAATCAAGCCAGTGCAGCACTGCAGTGC 1101
DB 959 -----AAGCGTGGCTTCGCGCAAGCGATTTCACTGCGGCGCATCGATCGAATGC 1014
QY 1102 CGCATCACGGAAGATCCAAACAAAGCTTCGCGCCAGATACGGAATCATCACCGG 1161
DB 1015 CGTATCAATGCGGAAGATCCGGAATACAAATTTCCGCGCAATCTTGGCGCATTAACAGC 1074
QY 1162 TACCGCTCACAGGCGGAGTGGCTGCTTGAAGTGGTGAAGTGAAGTGGTGGTTCGCA 1221
DB 1075 TATTTACGCGCGGCGCGCGCTGCTGCTGCAATTTTATACCGCACTACCGAA 1134
QY 1222 ATCAGCGCACATTTGACTTCAATGCTGTGTAAGTGAAGTGAAGTGGTGGTTCGCA 1281
DB 1135 ATTCGCGCTATTAGCATTCGCTGATTCGCAATTTGATTTGCTGGGGTGCAACGCGAA 1194
QY 1282 ACTGTGTGTGCTGCGACAGCGCGCTTGGCTGAGTTTCAACGCTGTCTGGTGTTCGAAC 1341
DB 1195 GAGGCGATCGCGCGATGCGCGTCTCTGCGGAAATGCGCATCACCGGCTTGGCGAG 1254
QY 1342 AACATTGGTTTCTTGTGCTGCTGCGGAAAGAGGACTTCACTTCCAAGCGCATCGCC 1401
DB 1255 ACCCTAGTTTCCATCAGCTGATTTGCAGATGCTGAGTTCTCTGCGGGGAACCTCTAT 1314
QY 1402 ACGGATTTATCG 1414
DB 1315 ACCAATTTGTTG 1327

RESULT 8
US-09-433-043B-5
; Sequence 5, Application US/09433043B
; Patent No. 6399342
; GENERAL INFORMATION:
; APPLICANT: HASELKORN, ROBERT
; APPLICANT: GORNICKI, PIOTR
; TITLE OF INVENTION: CYANOBACTERIAL AND PLANT ACETYL-CoA CARBOXYLASE
; FILE REFERENCE: ARCD:338US
; CURRENT APPLICATION NUMBER: US/09/433,043B
; PRIOR FILING DATE: 1999-10-25
; PRIOR FILING DATE: 1995-06-07
; PRIOR FILING DATE: 1995-06-07
; PRIOR FILING DATE: 1992-10-02
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 5
; LENGTH: 1362

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-433-043B-5

Query Match 8.8%; Score 306.6; DB 4; Length 1362;
Best Local Similarity 53.6%; Pred. No. 2.3e-75;
Matches 714; Conservative 0; Mismatches 604; Indels 15; Gaps 3;
QY 85 TTCAAAAGATCTTGGTAGCAAAACCGCGGAAATCGCGGTTCGCTTCCTGTCAGCA 144
DB 7 TTCAACAGATCTTGGTAGCAAAACCGCGGAAATCGCGGTTCGCTTCCTGTCAGCA 66
QY 145 CTCGAAACCGGTGAGCCACGAGTGTATTTACCCCGTGAAGATCGGGGATCTTCCAC 204
DB 67 CAAGAATCTGGGATCGGACGATCGCGGTTCACCTCCACTGTGGATCGCAACGCGCTCAT 126
QY 205 CGCTCTTTTGGCTTCTGAAGCTGTCCGATTTGGTACTGAAGCTCACAGTCAAGCGGTAC 264
DB 127 GTGCAGTTAGCGGACGAAGCGGTCTGTATTGG---CGAAGCGGCGAGCAGCAAAAGCTAT 183
QY 265 CTGACATCGATGAATATCGGTGACGTAAAGAGTTAAAGCAGATGCTATTACCCG 324
DB 184 TCAATATCCCAACATCATTTGCGGCGGCTGACCCCTAATGTCAGCGCCATTACACCC 243
QY 325 GGATATGGCTTCTGTCTGAAATGCCAGCTTGCCTCGAGTGCAGGAGGAAACGGCAT 384
DB 244 GGCTATGGCTTCTTGGCGGAGATGCCGCTTTGCAGAAATCGCGCGATCACCATCTC 303
QY 385 ACTTTATTTGGCCCAACCCAGAGGTTTGTATCTACCGGTGATAGTCTCGTGC---G 441
DB 304 ACCTTTATTTGGCCCAACCCAGGCTTTCGAGGCTGAGGCGATGAGTATCCACGCTAAG 363
QY 442 GTAACCGCGGAGAGGCTGTCTGCCAGTTTGGCGGATCCACCCGAGCAAAAC 501
DB 364 GAACAATGAGCGGTGCGGCTTCCGAGATTCGGGCACTGAGTCTGCTGACGGAT 423
QY 502 ATCGATGACATCGTTTAAAGCGCTGAAGCGGACAGTATACCCCATCTTTTGAAGCGCAT 561
DB 424 GTTGATTCGGCTGCCAAGTTGCTGCCAGATCGGCTATCCCGTCATGATCAAGCGACG 483
QY 562 CCGGTGTGCGGACGCGGTATGGCTTTGTTTCTTCACTGATGAGTCCGCAAAATG 621
DB 484 CGGGGGCGGTGTGCGGATGCGGTGTGGTGTGAGTCCCTGCGAGATCTGGAAAACTG 543
QY 622 GCAACAGAGCATCTCGTGAAGCTGAAGCGCATTCGCGAGGTTCCGTTATATGCGAA 681
DB 544 TTCCTGCTGCCAAGGAGAGCGGAGCGAGCTTTTGGGAATCCAGGACTGTATCTCGAA 603
QY 682 CGTGTGTGATTAACCCCGACACATTTGAAGTCAGATCTTGGCGATCGCACTGGAGAA 741
DB 604 AAATTTATCGATCGCCAGCGCATGTTGAATTTAGATCTTGGCGATGCTACGGCAAT 663
QY 742 GTTGACACCTTTATGAAGTGTGCTCACTGAGCGTGTCTACCAAAAAGTTGTGAA 801
DB 664 GTAGTGTATCTAGCGAGCGGATTTGCTCCATTCAACTGCTCCACCAAAAGTGTGAA 723
QY 802 ATTGCGGACGACAGCATTTGGATCCAGACTGCTGATTCGCAATTTGCGGATGCGAGTA 861
DB 724 GAAGCCCGCTAGTGGCGCTATCGCGAGACTGCGGCAAAAATGGGCGATGCGCGCGTC 783
QY 862 AAGTTCTGCCGCTCCATTGGTTTACAGGGCGGGAAACCGGTGAATCTTGGTGCATGAA 921
DB 784 AAGTCGCTCAAGCATCGGCTCATCGGTGCCGACCGTGGAGTTTCTGTCGATGCG 843
QY 922 AAGGCAACCACTTTTCAATGAATGAACCCAGTATCCAGGTGTGACACACCGTCACT 981
DB 844 ACCGGCAACTTCTACTTCATGGAGATGAATACCGGCTATCCAGTCCAGTCCAGTACA 903
QY 982 GAAGAGTCAACGAGTGGAGCTGTGAGCGGAGATGCGCTTGGCTGTGGTGGCAACC 1041
DB 904 GAAATGATTACGGGACTGGACTTGAATTGCGGAGCAGATTCGGATTGCGCCAAAGCGG----- 958

QY 1042 TTGAAGAAATGGGTCTGACCCCAAGATAAGATCAAGACCACCGGTGACGACTCAGTGC 1101
DB 959 ----AAGCGCTGCGCTTCCGCAAGCGGATATTCAACTGCGCGCCATGCGATCGAATGC 1014
QY 1102 CGCATCACCAAGGAGATCCAAACAACGGGTTCCGCCAGATACCGGAATCTATCACCGG 1161
DB 1015 CGTATCAATCGGAAGATCGGAATACAATTTCCGCCGAATCTTCCGCCATTTACAGG 1074
QY 1162 TACCGCTCACCAAGCGGAGTGGGTTTCTGCTTTGACGGTGCAGCTCGGTGGGAA 1221
DB 1075 TATTACCGCGCGGCGCGGCTTCTGCTGATTCCTCCATGTTTATACGACTACGAA 1134
QY 1222 ATCAACCCACACTTTGATCATGCTCGTGAATGAAATGACCTGCGGTGTTCCGATTGAA 1281
DB 1135 ATTCCGCTTATTACGATTCGCTGATTGGCAATGATTGCTGGGTGCAACACGGGAA 1194
QY 1282 ACTGCTTCTGCTGTCAGACGCGGCTTGGCTGAGTTCACCGTGTCTGGTGTGCAACC 1341
DB 1195 GAGCGGATCGCGGATGTCAGCGTGTCTTCCGGAATGCGCCATCACCGGCTTCCGACG 1254
QY 1342 AACATGCTTCTTTCGCTGCTGCTGCGGGAAGAGGACTTCACTTCAAGCGCATCGCC 1401
DB 1255 ACCCTTATTTCCATCAGCTGATGTTGCAGATGCTGCTGCGGGAATGCGCCATCTAT 1314
QY 1402 ACCGATTTATCG 1414
DB 1315 ACCAACTTTGTTG 1327
RESULT 9
US-08-611-107-7
; Sequence 7, Application US/08611107
; Patent No. 5801233
; GENERAL INFORMATION:
; APPLICANT: Haselkorn, Robert
; TITLE OF INVENTION: NUCLEIC ACID COMPOSITIONS ENCODING
; TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE AND USES
; TITLE OF INVENTION: THEREFOR
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,107
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US SN 07/956,700
; FILING DATE: 02-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US93/09340
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US SN 08/422,560
; FILING DATE: 14-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: ARCD:221
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1362 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-611-107-7

Query Match 8.8%; Score 305; DB 1; Length 1362;

Best Local Similarity 53.5%; Pred. No. 6.4e-75;

Matches 713; Conservative 0; Mismatches 605; Indels 15; Gaps 3;

QY 85 TTCAAAAAGATCTTTGGTAGCAAAACCGCGGGAATCCGGTCCGGTGTCTTTCCGTGACGCA 144
DB 7 TTCAACAAGATCTCTGATCGCCAAATCGCGGGAATCCCGCTTGGCATTTCTCCGCACTTGT 66
QY 145 CTCGAACCCGGTCAGCCACGATGCTATTATACCCCGTGAAGATCGGGATCATTTCCAC 204
DB 67 GAAGAACTCGGGATCGGCAGCATCGCGTTTCACTCCACTGTGGATCGCAACGCGCTCCAT 126
QY 205 CGCTCTTTTGGCTCTGAGCTGTCGATGTTGTAAGTCTCAAGGCTCACCGTCAAGGGGTAC 264
DB 127 GTGCACTTAGCGGAGAGCGGCTGTGATTGG---CGAAGGCGCAGCAGCAAAAGCTAT 183
QY 265 CTGACATCGAATTAATATCGGCGAGCTAAAAGATTAAAGCAGATGCTATTATCCCG 324
DB 184 CTCATATCCCAACATCATTTGCGGCGCCTGACCGTATGCGCGCATTCACCCC 243
QY 325 GGATATGGCTTCTGTCTGAAAATGCGGCTTCCCGCGAGTGGCGGGAAGCGGATTT 384
DB 244 GGCTATGGCTTCTTTGGCGGAGAAATCGCGCTTTTGCAGAAATCTCGCGCGATCACCATCTC 303
QY 385 ACTTTTATGTCACCAACCCAGAGGTTCTTGATCTCAACCGGTGATAAGTCTGTGTC---G 441
DB 304 ACTTTTATGTCACCAACCCAGAGGTTCTTGATCTCAACCGGTGATAAGTCTGTGTC---G 363
QY 442 GTAAACCGCGCGAAGAGGCTGTGTCAGCTTTTGGCGGAATCCACCCGAGCAAAAC 501
DB 364 GAAACAATGCAAGCGGCTCGCGTTTCCGACGATTCGGGCGAGTGAAGTCTGCTGACGAT 423
QY 502 ATCGATGACATGTTTAAAGCGCTGAAAGCGCAGACTTACCCCATCTTTTAAAGGCGAGTT 561
DB 424 GTTGATTCGGCTGCCAAGTTGCTGCCAGATCGGCTATCCCGTCATGATCAAAAGCAGC 483
QY 562 CGCGTGTGGCGGAGCGGATGCGTTTCTTCTTCCCTGATGAGCTCGGCAAAATG 621
DB 484 CGCGGGGCGGTTGTCGGGATGCGGCTGTGCGTGAAGCTCGAATCTGGAATACTG 543
QY 622 GCAACAGAAGCATCTCTGAAGCTGAAGCGCATTCGCGAGCGTTCCGTATATGTCGAA 681
DB 544 TTCCTTGTGTCGCAAGGAGAGCGGAGCAGCTTTTGGAAATCCAGACTGTATCTCGAA 603
QY 682 CGTGTGTGATTAACCCCGACACATTAAGTGCAGATCTTTGGCGATCGCACTGGAGAA 741
DB 604 AAATTTATCGATCGCCCGACGCGCATTTGAAATTTTCAAGTCTTTGGCGGATGCGTACGCAAT 663
QY 742 GTTGTACACTTTTATGAACGTGACTGCTCACTGCGAGCGTCTGCAACAAAAGTTGTGAA 801
DB 664 GTAGTGCATCTAGGGGAGCGGATTTGCTTCCATTTCAACGTGTCACCAAAAGCTGCTCGAA 723
QY 802 ATTGCGGCAGACAGCATTTGATCCAGAACTCGGTGATGCGATTTGTGCGGATGCGAGTA 861
DB 724 GAAGCCCCCAGTCCGCGCTATCGGCAGACCTTCGCGCAGAAAATGGGCGATGCGCGGTC 783
QY 862 AGTTCTGCGCTTCCATTTGTTTACAGGCGCGGGAACCGTGAATTTCTTGGTGCATGAA 921
DB 784 AAGTCTGCTCAAGCATCGGCTACATCGTGTGCGGCGACCGTGGAGTTTCTTGTGATGCG 843
QY 922 AAGGCGAACCATGTTTTCATGAATGAACCACTATTCAGGTTGAGCAGCACCGTCACT 981
DB 844 ACCGCAACTTCTACTTTATGAGATGAATACCGCATCCAGTCGAGCATCCAGTCACA 903
QY 982 GAAGAAGTCCACGAGTGACCTGTGTAAGCGGCAGATCGGCTTGGCTGTGTTGTCGACAC 1041

DB 904 GAAATGATTACGGACTTGGACTTGTGATTCGGAGCAGATTCGGATTGCCAAGGCG----- 958
QY 1042 TTGAAGGAATTTGGTCTTGACCCAAAGATAAGATCAAGACCCACGGTGCAGCACTCACTGTC 1101
DB 959 ---ANGCGCTGCGCTTCCGCAAGCGATATTCAACTGCGCGGCCATCGATCGAATGC 1014
QY 1102 CGCATCACCGAGAGATCCAAACAAAGGCTTCCGCCAGATACCGGAATATATCACCGCG 1161
DB 1015 CATAATATCGGAGATCCGGAATACAATTTCCGCGGATCTTGGCGCATTTACAGGC 1074
QY 1162 TACCGCTCACAGCGGAGCTGCGTCTTGAACGGTGCAGCTCAGCTCGGTGGGAA 1221
DB 1075 TATTTACCGCGCGCGCGCTTGTGTCGATTTCCCATGTTTATACGACTACGAA 1134
QY 1222 ATCACCGCACATTTGACTCCATGCTGTAAGTGAATGACCTGCGGTTCGACTTTGAA 1281
DB 1135 ATTCCGCTTATTAGATTCGCTGATTTGGCAATTTGATTGTCTGGGTGCAACACGGGAA 1194
QY 1282 ACTGCTGTTGCTGTCAGCAGCGCGGCTTGGCTGAGTTTCACTGTTCTGTTGTCGAA 1341
DB 1195 GAGGCGATCGCGCGGATGCGCGTCTCTCGGGAATGCGGCATCACCGGCTTGCAGC 1254
QY 1342 AACATTGTTTCTTGGTGTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCC 1401
DB 1255 ACCCTTAGTTTCCATCAGCTGATGTTGCAGATGCTGTGATTTCTGCGGCGGAACTTAT 1314
QY 1402 ACCGGATTATCG 1414
DB 1315 ACCAATTTGTTG 1327

RESULT 10

US-08-422-560A-7

Sequence 7, Application US/08422560A

Patent No. 5910626

GENERAL INFORMATION:

APPLICANT: Haselkorn, Robert

APPLICANT: Gornicki, Piotr

TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND

TITLE OF INVENTION: METHODS FOR USE

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: TX

COUNTRY: USA

ZIP: 77210-4433

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/08/422,560A

FILING DATE: 14-APR-1995

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/956,700

FILING DATE: 02-OCT-1992

ATTORNEY/AGENT INFORMATION:

NAME: Wilson, Mark B.

REGISTRATION NUMBER: 37,259

REFERENCE/DOCKET NUMBER: ARCD:152/WIM

TELECOMMUNICATION INFORMATION:

TELEPHONE: 512-418-3000

TELEFAX: 512-474-7577

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 1362 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear
US-08-422-560A-7

Query Match 8.8%; Score 305; DB 2; Length 1362;
Best Local Similarity 53.5%; Pred. No. 6.4e-75;
Matches 713; Conservative 0; Mismatches 605; Indels 15; Gaps 3;

QY 85 TTCAAAAGATCTTGGTAGCAAAACCGGGCGAAATCGGGTCCGTTCCTCGTGACGA 144
DB 7 TTCACAAAGATCTTGGTAGCAAAACCGGGCGAAATCGGGTCCGTTCCTCGTGACGA 66

QY 145 CTCGAAACCGGTGACGACCGAGTGTATTTACCCCGTGAAGATCGGGATCATTCAC 204
DB 67 GAAGACTCGGGATCGGACGATCGCGTTCACCTGCTGGAATCGAAGCGCTCAT 126

QY 205 CGCTCTTTTGTCTGAAGCTGTGCGATTTGGTACTGAAGGCTCACCAGTCAAGCGGTAC 264
DB 127 GTGCAGTTAGCGGACGAAGCGGTCTGTATTGG---CGAAGCGCGCAGCAGCAAAAGCTAT 183

QY 265 CTGACATCGATGAATATTCGGTGCAGCTAAAGAGTAAAGCAGATGCTATTATTCGG 324
DB 184 CTCAATATCCCAACATCATTTGCGGCGCCCTGACCCGTATGTCAGCGCCATTCACCCC 243

QY 325 GGATATGCTTCTGCTGTAAGATGCCAGCTTCCCGCGAGTCCCGGAAACCGCAT 384
DB 244 GGCTATGGCTTCTTGGCGGAGATGCCCGCTTTGAGAAATCTGCGCGATCACAATCTC 303

QY 385 ACTTTTATGCGCCAAACCCAGAGTCTTCTGATCTCACCGGTGATAGTCTCTGTC--G 441
DB 304 ACTTTTATGCGCCAGCCCGGATTCGATTCAGCCATGGCGATNAATCCACGCTAAG 363

QY 442 GTAACCGCGCGAAGAGGTGTGTGTCAGTCTTGGCGGAATCCACCGGCAAAAC 501
DB 364 GAAACAAATGACAGCGGTGCGGCTTCCGACGATTCGCGCAGTGCAGTCTGCTGACGAT 423

QY 502 ATCGATGACATCGTTAAAGCGCTGAAGCCAGACTTACCCCATCTTTGTAAGCGAGTT 561
DB 424 GTTGAITCGGCTGCCAAGTTGCTGCGCGAGATCGGCTATCCCGTCATGATCAAGCGACG 483

QY 562 GCCGCTGTGCGGACCGGTATGCGTCTTCTTCCATCGATGAGTCCGCAAAATG 621
DB 484 GCGGGGCGGTGTGCGGTATGCGGTGTGCGTGTGAGCTTCGAGATCTGGAATACTG 543

QY 622 GCAACAGAGCATCTGTGAAGCTGAAGCGGATTCGCGGAGTTCGGTATATGTGAA 681
DB 544 TTCCTGTGTCGCGGAGGAGCGGAGCGAGTCTTGGGAATCCAGGACTGTATCTCGAA 603

QY 682 CGTCTGTGATTAACCCCGACCATTTGAAGTGCAGATCTTGGCGATCGCACTGGAGAA 741
DB 604 AAATTTATCGATCGCCCAACGCGCATTTGAATTTTCAGATCTTGGCGGATGCTACGCGAAT 663

QY 742 GTTGTACACCTTTATGAAGCTGACTGTCTCACTGACGCGTCTGACCAAAAGTTGTGAA 801
DB 664 GTAGTGATCTAGGAGCGGATGTCTCAATCAAGCTGTCAAAAGCTGCTCGAA 723

QY 802 ATTGCGCGACGACGATTTGGATCCAGAACTCGGTGTATGCGATTTGCGGATGCACTA 861
DB 724 GAAGCCCCAGTCCGGGCTATCGGACAGCTGCGGAGAAATGGGCGATGCGCGCGTC 783

QY 862 AAGTTCTGCGCTCATTTGTTACAGGGCGGGAGCCGTTGGTGAATCTTGGTGTGATGAA 921
DB 784 AAAGTCTCAAGCGATCGGCTATCATCGTGTGCGGACCGGTGAGATTTCTGTCGATGCG 843

QY 922 AAGGGCAACACGCTTTTTCATCGAAATGAACCGATGATCCAGGTTGAGCACACCGTGA 981
DB 844 ACCGGCAACTTCTACTCATGAGATGAATACCGCATTCAGTTCGAGATCCAGTACA 903

QY 982 GAAGAAGTCAACGAGGTGACCTGGTGAAGCGGAGATGCGCTTGGTGTGTTGCAACC 1041
DB 904 GAAATGATTACGGGACTGGAATGATTGCGGAGCAGATTCGATTGCGGAGGCG----- 958

QY 1042 TTGAAGGAATTTGGTCTGACCCCAAGATGAATCAAGCCCAAGGTCGAGCTGAGTGC 1101

DB 959 ----AAGCGTGGCTTCCGGCAAGCGGATATTCAACTGCGCGCCATGCGATGCAATGC 1014

QY 1102 CGCATCCACGGAAGATCAAAACAGCGGTTCGCGCCAGATACCGAACTATCACCGCG 1161

DB 1015 CGTATCAATCGGAAGATCGGAATACAATTTCCGGCGGAATCTCGCGCATTTACAGGC 1074

QY 1162 TACCGCTCACAGCGGAGTGGGTTCTGCTTGAACGGTGCAGCTCAGCTCGGTGGGAA 1221

DB 1075 TATTTACCGCGCGCGCGCGGCTTGGTGTGATTTCCCATGTTTATACGACTACGAA 1134

QY 1222 ATCACCCACACACTTTGACTCCATGCTCGTGAATAATGACCTGCCGTGGTTCGCACTTTGAA 1281

DB 1135 ATTCCGCGCTATTACGATTCGCTGATTTGGCAATGATTTCTGGGTGCAACAGCGAA 1194

QY 1282 ACTGCTGTTGCTGTCACAGCGCGCTTGGCTGAGTTCAACCGTGTCTGTTGTTGCAAC 1341

DB 1195 GAGCGATCGCGGATGACAGCGTCTCTCGGGAATGCGCATCACCGGCTTCCGACG 1254

QY 1342 AACATTTGTTCTTGGTGGTTCCTGCGGGAAGAGAGACTTCACTTCAAGCGCATCGCC 1401

DB 1255 ACCCTTAGTTTCCATCAGTGTGTCAGATGCTGCTGAGTCTCTGCGCGGAACTCTAT 1314

QY 1402 ACCGATTTATCG 1414

DB 1315 ACCAACTTTGTTG 1327

RESULT 11
US-08-468-793-7
; Sequence 7, Application US/08468793
; Patent No. 6177267
; GENERAL INFORMATION:
; APPLICANT: Haselkorn, Robert
; APPLICANT: Gornicki, Piotr
; TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,793
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/422,560
; FILING DATE: 14-APR-1995
; APPLICATION NUMBER: US SN 07/956,700
; FILING DATE: 02-OCT-1992
; CLASSIFICATION: 800
; APPLICATION NUMBER: PCT/US93/09340
; FILING DATE: 30-SEP-1993
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: ARCD:152/KIT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1362 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
US-08-468-793-7

Query Match 8.8%; Score 305; DB 4; Length 1362;

Best Local Similarity 53.5%; Pred. No. 6.4e-75;

Matches 713; Conservative 0; Mismatches 605; Indels 15; Gaps 3;

Qy 85 TTCAAAAGATCTTTGGTAGCAAAACCGCGGAAATCGCGTTCGCTTCCTCGTAGCA 144
Db 7 TTCAAAAGATCTTTGGTAGCAAAACCGCGGAAATCGCGTTCGCTTCCTCGTAGCA 66
Qy 145 CTGAAACCGGTGCGAGCAGGTAGCTATTATACCCCGTGAAGATCGGGGATCATTTCCAC 204
Db 67 GAAGAACTCGGGATCGGCAGCATCGCGTTTCACTCCACTGTGATGCAACGCGTCCAT 126
Qy 205 CGCTCTTTTGTCTGTAAGCTGTCCGATTTGGTACTGAAGGCTCACAGTCAAGGGGTAC 264
Db 127 GTGCAGTTAGCGGACGAGCGGTCTGTATTGG---CGAAGCGCCAGCAGCAAAAGCTAT 183
Qy 265 CTGACATCGATGAATATCGGTGACGCTAAAGAGTTAAAGCAGATGCTATTACCCG 324
Db 184 CTCAATTTCCCAACATCAITTCGCGGCTGACCGTAAATGCCAGCGCATTCACCCC 243
Qy 325 GGATATGGCTTCCTGTGAAATGCCAGCTTCCCGCGAGTCCCGGGAACGCAIT 384
Db 244 GGCATATGCTTCTTTGGGAGATGCCCGCTTTCGAGAAATCTGCGCGGATCACCATTC 303
Qy 385 ACTTTATTTGCCCCAACCCAGAGGTTCTTGATCTCACCGGTGATTAAGTCTCGTGC---G 441
Db 304 ACTTTATTTGCCCCAGCCCGGATTCGATTCGAGCCATGGCGGATTAATCCACCGTAA 363
Qy 442 GTAACCGCGCGAAGAGGCTGTCTGCGATTTTGGCGGAATCCACCCGAGCAAAAC 501
Db 364 GAAACAAATGACAGCGGTGCGGCTTCCGACGATTCGCGGACGTCTGTGAGCGAT 423
Qy 502 ATCATGACATCTGTTAAAGGCTTGAAGGCGAGACTTACCCCATCTTTTGAAGGAGTT 561
Db 424 GTTATTCGGCTGCCAAGTTGCTGCCAGATCGGCTATCCCGTCATGATCAAAAGCA 483
Qy 562 GCGGTGTGCGGACGCGGTATCGCTTTGTTTCTTCACTGATGAGTCCGCAATTTG 621
Db 484 GCGGGGCGGTGTGCGGTATGCGGTGTGCGGTGAGCTGAGCTCTGGAACAC 543
Qy 622 GCAACAGAGCATCTCGTAGCTGAGCGGATTCGCGGACGTTCCGTTATATGCGAA 681
Db 544 TTCCTTCTGCTCCAGGAGAGCGGACGACTTTTGGGATCCAGGACTGTATCTCGAA 603
Qy 682 CGTGTGTGATTAACCCCGACACATTGAATGAGATCCTTGGCGATCGACTGGAGAA 741
Db 604 AAATTTATCGATCGCCCGACGCGTGAATTTGAGATCTTGGCGGATGCTACGGCAAT 663
Qy 742 GTTGACACCTTTATGAACGTGACTGCTCACTGAGCGGTGCTACCAAAAAGTTGTGAA 801
Db 664 GTAGTGATCTAGGCGAGCGGATTTGCTTCAATTCAGCTGCTACCAAAAAGCTGCTGAA 723
Qy 802 ATTGCGGACAGACAGCATTTGGATCCAGACTGCGGTGATCGCAITTTGCGGATGCA 861
Db 724 GAAAGCCCGAGTCCGCGCTATCGGACACCTGCGGAGAAAATGGGCGATGCGCGCTC 783
Qy 862 AAGTTCTCGCTTCAATTTGTTTACAGGCGCGGAAACCGTGGAAATCTTGGTTCGATGAA 921
Db 784 AAGTCTCGTCAAGCGATCGCTACATCGTCCGCGACCGTGGAGTTTCTGGTCTGATGCG 843
Qy 922 AAGGCGACACCGTTTTCATGAATGAACACGATTCAGGTTGAGCAGACCGTCACT 981
Db 844 ACCGCAACTTCTACTTCAATGAGATGATACCCGCTCCAGTTCGAGTCCAGTCCACA 903
Qy 982 GAAGAGTCCAGGCTGAGCTGTTGAAGGCGGAGATGCGTGGCTGCTGGTGCAC 1041
Db 904 GAATGATTAACGGGACTGACTTGTGATTCGAGAGCAGATTCGATTCGCAAGGCG----- 958

Qy 1042 TTGAAGGAATTTGGTCTGACCCCAAGATAAGATCAAGACCCACCGTGCAGCACTGCAGTGC 1101
Db 959 ----AAGCGCTGCGCTTCCGSCAAGCGATATTCACTGCGCGCCATCGATCGAATGC 1014
Qy 1102 CGCATCACCGAGGATCCAAACAGCGGCTTCCGCCAGATACCGGAACATATCACCGCG 1161
Db 1015 CGTATCATGCGGAGATCCGGAATCAATTTCCGCGGATCTCTGCCCAITACAGGC 1074
Qy 1162 TACCGCTCACCGAGGAGTGGCGTTTCTTTCAGCGTGCAGTGCAGTGGTGGGAA 1221
Db 1075 TATTTACCGCGCGCGCGCGGCTTCTGTCGATTCCTCCATGTTTATACGACTACGA 1134
Qy 1222 ATCACCGCACATTTGACTCCATGCTGCTGAAATGACCTGCGCTGTTTCGACTTTGAA 1281
Db 1135 ATTCCGCCCTATTACGATTCGCTGATTGGCAATTTGATGTTCTGGGTGCAACCGGAA 1194
Qy 1282 ACTGCTGTGCTGTCGACAGCGCGGTTGGCTGAGTTCACTGCTGTTGTTGCAACC 1341
Db 1195 GAGCGCATCGCGCGGATGCGAGCGTCTTCCGGAATGCGCATCACCGGCTTCCCGACG 1254
Qy 1342 AACATTGTTTCTTGGTGGCTTCTGCGGAGAGGAGTTCCTTCAAGCGCATCGCC 1401
Db 1255 ACCCTTAGTTTCCATCAGCTGATGTTGAGATGCTGAGTTCCTGCGCGGGAACCTCTAT 1314
Qy 1402 ACCGATTTATCG 1414
Db 1315 ACCAATTTGTTG 1327

RESULT 12

US-08-074-121-1

Sequence 1, Application US/08074121

Patent No. 5767362

GENERAL INFORMATION:

APPLICANT: Best, Elaine

APPLICANT: Knauf, Vic C.

TITLE OF INVENTION: Methods and Compositions for Modulating

TITLE OF INVENTION: Lipid Content of Plant Tissues

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson

STREET: 2200 Sand Hill Road, Suite 100

CITY: Menlo Park

STATE: California

COUNTRY: US

ZIP: 94025

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA: US/08/074,121

FILING DATE: 08-JUN-1993

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Rae-Venter, Barbara

REGISTRATION NUMBER: 32,750

REFERENCE/DOCKET NUMBER: 05938/043001

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 854-5277

TELEFAX: (415) 854-0875

TELEX: 200154

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3077 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 800..1267
FEATURE: CDS
NAME/KEY: 1281...2626
LOCATION: 1281...2626
US-08-074-121-1

Query Match 8.78; Score 302.2; DB 1; Length 3077;
Best Local Similarity 54.9; Pred. No. 66-74;
Matches 714; Conservative 0; Mismatches 563; Indels 24; Gaps 5;

Qy	89	AAAGATCTTGGT	AGCAAAACCGCGCGGAAATCGCGGTCCGCTTCGTCGACGACTCG	148
Db	1288	ATAAAATGTTAT	TGCCAACCGCGCGGAGATTGATTCGCTATCTTCGTCGCTGTAAG	1347
Qy	149	AAACCGGTGAC	CCACGGTAGCTATTTACCCCGTGAAGATCGGGGATCATTCACCGCT	208
Db	1348	AACCTGGGCAT	CAAGACTGTGCTGTGCACTCCAGCGGGATCGCGATCTAAACACGAT	1407
Qy	209	CTTTTGCTTCT	GAGCTGTCGCAATGGTACTGAAGCTCACAGTCAAGCGGTACCTGG	268
Db	1408	TACTGGCAGAT	GAACCGTCTGTATGGCCCTGCTCCGT---CAGTAAAGTTTATCTGA	1464
Qy	269	ACATCGATGAA	ATATCGGTGACGTAATAAAGTTAAAGCAGATGCTATTTACCCGGAT	328
Db	1465	ACATCCGGCAT	ATCATAGCCCGCTGAATCACCGCGGAGTAGCAATCCATCCGGTT	1524
Qy	329	ATGCTTCTCTG	TCGAAATGCCAGCTTGCCCGGAGTGC CGGAAACGGCAATTACTT	388
Db	1525	ACGGTTCCTCT	CGGAAACGCCAACTTTGCCGAGCAGGTTGAACGCTCCGGCTTTATCT	1584
Qy	389	TTATTGGCCCA	CCCCAGAGGTTCTTCACTCACCGGTGATAAGTCTCGTGGTAACCG	448
Db	1585	TCATTGGCCCC	GAAGCAGAAACCAATTCGCCTGATGGCGGCAAAATATCGCAATCCGG	1644
Qy	449	CCGCAAGAAAG	CGTGCTGCGAGTTTGGCGGAATC-----CACCCGAGCAAAACA	502
Db	1645	CGATGAATAAA	AGCGGGTCCCTTCGTACCGGGTTCTGACGGCCGCTGGCGACATA	1704
Qy	503	TCGATGACATG	TTAAAGCGCTGAAGCCAGACTTACCCCATCTTTGTAAGCAGTTG	562
Db	1705	TGGATAAAAC	CCGTGCCATTGCTAAACGCAATGCTTCGGTGATTATCAAGCCCTCG	1764
Qy	563	CCGGTGTGGCG	AGCGGTATGCGCTTTGTTTTCACCTGATGAGTCCGCAAAATGG	622
Db	1765	CGCGCGCGCGG	TCGCGGTATGCGCGTAGTGGCGCGAGCTGGAATCGCAATCCA	1824
Qy	623	CAACAGAAGCA	TCTCGTGAAGCTCAAGCGCATTCGCGCAGGTTTCGCTATATGTCGAAC	682
Db	1825	TCTCCATGAC	CCCGTGGAGCGAAGAGCTGCTTTCAGCAACGATATGTTTACATGAGA	1884
Qy	683	GTGCTGTGATT	AAACCCCGACACATTGAAGTGCAGATCTTTGGCGATCGCACTGGAGAAG	742
Db	1885	AATACCTGGA	AAATCTTCGCGAGTTCAGGTACTGGCTGACGGTCAGGCAACG	1944
Qy	743	TTGTACACCTT	TATGAACGTGACTGCTCACTGAGCGTGCACCAAAAGTTGCGAAA	802
Db	1945	CTATCTATCT	TCTGGCGGAACGTAAGTCTGCTCAATGCAACCGCGCACAGAAAGTGTGGAAG	2004
Qy	803	TTGGCGCAGCA	GACATTTGATCCAGAACTGCGTGATCGCATTTGTCGGATGCACTAA	862
Db	2005	AAGCGCCAGAC	CCGGGCAATTTACCCCGAATCGCGTCTGCTATCGCGCAAGTTGCGCTA	2064
Qy	863	AGTTCTCGCGT	CCATTTGGTTTACAGGCGCGGAAACCGTGAATTTCTTGGTCATGAAA	922
Db	2065	AAGCGTGTGTT	GTATATCGCGTATCGCGTGCAGGTACTTTTCGAGTCTCTGTTTC---GAAA	2121
Qy	923	AGGCAACCAAG	TTTTCATCGAAATGAACCACTATTCAGGTTGAGCAACCGGTGACTG	982
Db	2122	ACGCGAGTTCT	ATTTTCATCGAAATGAACCACTATTCAGGTTGAGCAACCGGTGACTG	2181
Qy	983	AAGAAGTACCG	AGGTCGACTGTTGAAGCGGAGATGCGCTTGGCTGCTGGTGCACTT	1042
Db	2182	AAATGATCAC	CGCGTTGACCTGTATCAAGAAACAGCTGCGTATCGCTGCGCGGTCAACCG	2241

RESULT 13

PCT-US94-06447-1
; Sequence 1. Application PC/TUS9406447
; GENERAL INFORMATION:
; APPLICANT: Calgene, Inc.
; TITLE OF INVENTION: Methods and Compositions for Modulating
; TITLE OF INVENTION: Lipid Content of Plant Tissues
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weil, Gotshal & Manges
; STREET: 2882 Sand Hill Road, Suite 280
; CITY: Menlo Park
; STATE: California
; COUNTRY: US
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/06447
; FILING DATE: 06-JUN-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rae-Venter, Barbara
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: CGNE-097/WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 926-6200
; TELEFAX: (415) 854-3713
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3077 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 800..1267
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1281..2626
; PCT-US94-06447-1

Query Match 8.7%; Score 302.2; DB 5; Length 3077;
Best Local Similarity 54.9%; Pred. No. 6e-74;
Matches 714; Conservative 0; Mismatches 563; Indels 24; Gaps 5;

QY 89 AAAAGATCTTGGTAGCAAAACCGGGGAGAAATCGCGTCCGTTCCGTTGCGAGCACTCG 148
Db 1288 ATAAATTTGTTATTTGCAACCGCGGAGATTGCANTTGGCTATTTCTGTCCTGTAAAG 1347
QY 149 AAACCGGTGAGCAGCAGTATTTTACCCCGTGAAGATCGGGATCATTTCCACCGCT 208
Db 1348 AACTGGSCATCAAGACTGTCTGTGCTGCACTCCAGCGGGATCGGATCTAAACACGTAT 1407
QY 209 CTTTTGGCTTGAAGCTGTCGATTTGTTACTGAAGCTCACCAGTCAAGGCTACTCGG 268
Db 1408 TACTGGCAGATGAACCGTCTGTTATTTGGCCCTGCTCCGT---CAGTAAAGATTATCGA 1464
QY 269 ACATCGATGAATATTCGTTGCGAGCTAAAGAGTTAAAGCAGATGCTATTTACCCGGAT 328
Db 1465 ACATCCCGCAATCATCAGCGCGCTGAATACCGCGCAGTAGCAATCCATCCCGGTT 1524
QY 329 ATGGCTTCTGTCTGAATAATGCCAGCTTGCCTCCGAGTGCCTGGGAAACAGGCAATTCTT 388
Db 1525 ACGGCTTCTCTCGGAACCGCAACTTTTCCGAGCAGGTTGAACGCTCCGGCTTTATCT 1584
QY 389 TTATTTGCCCAACCCCAAGAGGTTCTTGTATCTCACCGGTGATAAGTCTCTGTCGGTAAACG 448
Db 1585 TCATTTGCCGGAAGAGCAGAAACCAATTCGCTGTATGGGCGCAAAAGTATCCGCAATCGCG 1644
QY 449 CCGGAAGAGGCTGTCTGCAAGTCTTGGCGGAATC-----CACCCCGAGCAAAACA 502
Db 1645 CGATGAATAAAGCGGGGCTCCCTTGGCTACCGGTTCTGACGCGCCGCTGGCGCAGATA 1704
QY 503 TCATGATCATCTGTTAAAGCGCTGAAGCGCAGACTTACCCCATCTTTGTAAGGCGAGTTG 562
Db 1705 TGGATAAACCCTGCTCCATCTGTAACGCATTTGTTATCCGTTATATCAAGCCTCCG 1764
QY 563 CCGTGTGCGGACGCGGTATGCGTTCTTCTTTCACCTGATGAGTCCGCAATTCG 622
Db 1765 CGCGCGCGCGCTGCGGTATGCGGTATGCGGTAGTGGCGGACGCTGAACGTCACATCCA 1824
QY 623 CAACAGAAGCATCTGTTGAAGCTGAACCGGATTCGCGGACGTTGCGTATATGTCAAC 682
Db 1825 TCTCCATGACCGGTGCGGAAGCAAGCTCTTTACGAACAGATATGTTTACATGGAGA 1884
QY 683 GTGCTGTGATTAACCCCAAGCATTTGAAGTGCAGATCTTGGCGATTCGCACTGGAGAAG 742
Db 1885 AATACCTGGAAATCTCTCGCCAGCTCAGATTCAGGTACTGGCTGACGCTCAGGGCAACG 1944
QY 743 TTGTACACCTTTATGAACGTCAGTCTCACTGAGCGTCTGTCACCAAAAGTTGTGAAA 802
Db 1945 CTATCTATCTGGCGGAACGTCAGTCTCCATGCAACCGCGCCACCAAGAAAGTGTGGAAG 2004
QY 803 TTGCGCCAGCAGCATTTTGGATCCAGAACTGCGTGATCGCATTTTGGCGATGCGAGTAA 862
Db 2005 AAGCGCCAGCACCGGGCATTTACCCCGAATTCGCTGCTACATCGCGAAGCTTGGCTA 2064
QY 863 AGTTCTGCGCTCCATTGGTTTACCAGGCGCGGAAACCGTGAATCTTGTGATGATAA 922
Db 2065 AAGCGTGTGTATCGCTATCGCGTGCAGTACTTTCTGAGTCTCTGTC---GAAA 2121
QY 923 AGCGCAACAGCTTTTCATCGAATGAACCACTATCTCAGTTTGAAGCAGCAGTCACTG 982
Db 2122 ACGCGAGTCTATTTTATCGAATGAACCACTATCTCAGTTTGAAGCAGCAGTCACTG 2181
QY 983 AAGAAGTCAACGAGTGAACCTGTGAAGCGGAGATGCGTGTGCGTGTGCGTGTGCAACT 1042
Db 2182 AATGATCACCAGGCTGATCTGATCAAGAAACAGTGCCTGCTGCTGCGCGGTCACACCG 2241
QY 1043 TGAAGGAATTGGGTCTGACCCCAAGATAAGATCAAGACCCACGCTGCGTGAAGTCACTGCG 1102
Db 2242 TGTGCG-----ATCAAGCAAGAAGTTTACGTTTCGCGGCATTCGCGTGAATGTC 2292
QY 1103 GCATCACCAGGAAGATCCAAACAAACCGGCTTCGCCCGAGATACCGGAATCATCACCGCGT 1162

Db 2293 GTATCAACCGCGAAGATCCGAACA---CTTCTCCCAAGTCCGGCAAAATCACCGTT 2349
QY 1163 ACCGCTCAACAGGCGGAGCTGGCGTTGCTTGTGACGFTGCAAGCTCAGCTCGGTGGGAAA 1222
Db 2350 TCCAGGACCTTGGCGGTTTGGCGTACGTTGGGAGTCTCATATCTACGCGGCTACACCG 2409
QY 1223 TCACCGCACACTTTGACTTCCATGCTGTTGTAATGACCTGCGGTGGTTCCGACTTTGAAA 1282
Db 2410 TACCGCGTACTATGACTCAATGATCGTAACTGATTTGCTACGGTGAACACCGTGACG 2469
QY 1283 CTGCTGTTGCTGTCACAGCGCGCTTGGCTGAGTTGCTGAGTTGCTGCTGTTGCAACA 1342
Db 2470 TGGCGATTTCGCGATGAAGATGCGCTGCGAGGAGTGCATCATCGCGTATCAAAACCA 2529
QY 1343 ACATTGTTTCTTCCGTCGCTGCTGCGGGAAGAGGACTTC 1383
Db 2530 ACGTTGATCTGCAGATCCGATCATGATGACGAGAACTTC 2570

RESULT 14
US-07-956-700B-1
; Sequence 1, Application US/07956700B
; Patent No. 5539092
; GENERAL INFORMATION:
; APPLICANT: Robert Haselkorn and Piotr Gornicki
; TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
; NUMBER OF SEQUENCES: 116
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: 321 No. 5539092th Clark Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60610
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/956,700B
; FILING DATE: 19921002
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Thomas E. No. 5539092thrup
; REGISTRATION NUMBER: 33,268
; REFERENCE/DOCKET NUMBER: ARCD:058
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 1-312-744-0090
; TELEFAX: 1-312-755-4489
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3065 base pairs
; TYPE: Nucleic acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: Oligonucleotide
US-07-956-700B-1

Query Match 8.5%; Score 297; DB 1; Length 3065;
Best Local Similarity 54.4%; Pred. No. 1.7e-72;
Matches 691; Conservative 0; Mismatches 565; Indels 15; Gaps 4;

QY 85 TTCAAAAGATCTTGGTAGCAAAACCGCGGAGAAATCGCGTCCGTTCCGTCGACGA 144
Db 1314 TTGCAAAATATTAATTGCCATCGGGAGAAATAGCGTCCGATTTCTCCGCGCTGT 1373
QY 145 CTCGAAACCGGTGACGACCGTAGCTATTTACCCCGTGAAGATCGGGATCATTCAC 204
Db 1374 GAGGAATGGGAGTTGCGACGATCGAGTTATTTCGACTGTTGACCGGAATCTCTTAT 1433

```
QY 205 CGCTCTTTGCTTCTGAGCTGTCCGCAATTTGGTACTGAGGCTCACCAGTCAAGCGGTAC 264
Db 1434 GTCCAACTTGTCTGAGAGCGGTTGTATTGGC---GAACGTGTAGCGCTAANAAGTTAT 1490
QY 265 CTGACATCGATGAATATATCGGTGCGAGCTAAAAAGTTAAAGCAGATGCTATTACCCG 324
Db 1491 TTGAATATCCCAATATATTGCTGGGCTTTACGGCATGCCAGTGTCTTATCATCCT 1550
QY 325 GGATATGGCTTCTGCTGAAATATGCCAGCTTGCCTCCGCGAGTGGCGGAAAAACGGCAAT 384
Db 1551 GGGTATGGCTTTTATCTGAAATATGCCAAATTTGCGGAAATCTGTGCTGACCATCACAAT 1610
QY 385 ACTTTTATGGCCCAACCCACAGAGTTCTTGATCTCACCGGTGATAAGTCTCGTGGCGTA 444
Db 1611 GCATTCATTTGGCCCCCCCCCAGAGCTATCCGCTCATGGGGGCAAAATCCACTGCCAAG 1670
QY 445 ACCCGCGGAAGAAGGTGTCTGCCAGTTTGTGGCG--AATCCACCCCGAGCAAAAACA 502
Db 1671 GAACCATGCAAAAGCTGGTATCCGACAGTACGGGTAGTGAAGTTTGGTAGAGACA 1730
QY 503 TCGATGACATCGTTAAAGCGCTGAAGCCA-GACTTACCCCATCTTTGTAAAGCGAGTT 561
Db 1731 GAGCAAGAAGGATTAGAAGTCCGCAAGATATGGCTACCCAGTGATCAAAAGCCACG 1790
QY 562 GCCGGTGTGGCGACCGGCTATGGCTTTGTTTCTTCCACTGATGAGCTCCGCAATTTG 621
Db 1791 GCTGGTGTGGCGCGCGGGTATCGACTGGTGGCATGCGCCAGATGAATTTGTCAAACTG 1850
QY 622 GCAACAGAAGCATCTCGTGAAGCTGMAAGCGCATTCGCGACGCTCGGTATATGTCGAA 681
Db 1851 TTCTTAGCGCCCAAGGTGAAGCTTGGTGCAGCTTTGGTAATGCTGGCGTTTATAGAA 1910
QY 682 CGTCTGTGATTAACCCACGACATTAAGTGCAGATCTTTGGCGATCGCACTGGAGAA 741
Db 1911 AAATTTATGAAGCTCGCGCCACATTAATTTCAAAATTTTGGCTGATTAATACGGCAAT 1970
QY 742 GTTGTACACCTTTATGAAGTGTCTCTACTGAGCGTGTCTACCAAAAAGTTGTGAA 801
Db 1971 GTGATTCACTTGGGTGAGAGGGATTGTCTAAATCAGCGTCTGTAAACCAAAAGTTACTAGAA 2030
QY 802 ATTCCGCCAGCACAGCATTTGGATCCAGAACTGGTGTGATCGCATTTGTGGCGATGCACTA 861
Db 2031 GAAGCCCCAGCCAGCGTTGGACTCAGACTAAGGAGAAAAAATGGGACAAAGCGCGGTG 2090
QY 862 AAGTCTCCCGCTCCATTTGGTTACAGGGCGCGGAAACCGTGAATTTCTTGGTTCGATGAA 921
Db 2091 AAAGCGGCTCAGTTTATCAATTTACCGCGGGCAGGTACTATCGAGTTTTCGTAGATAGA 2150
QY 922 AAGGGCAACACAGTTTTCATCGAAATGAACCCAGTATCCAGGTTGAGCACACCGTGACT 981
Db 2151 TCCGGTCAGTTTACTTTTATGGAGATGAACACCCCGGATTCAGGTAGAACATCCCGTAAT 2210
QY 982 GAAGAAGTCAACGAGGTGACCTGGTGAAGCGCAGATGCGCTTGGCTGCTGGTGCACCC 1041
Db 2211 GAGATGTTTCTGGAGTGGATTATTTGGTTAGCAAAATCAGAAATGGCCCAAGGGGAAA-- 2268
QY 1042 TTGAAGGAATTTGGGTCTGACCCAAAGATAAGATCAAGACCCAGCTGCAGCACTCCAGTGC 1101
Db 2269 -----GACTTAGACTAATCTCAAGACCAAGTAGTTTATCCGGTCTATCGATCGAATGT 2321
QY 1102 CGATATCACCGGAAGATCCAAACAGAGGCTTCGCGCCAGATACCGGAATATCAACCGCG 1161
Db 2322 CGCATCAATGCGGAAGACCCAGACAGCATTTCCGCGCCAGCACCCCGGACGCAATTAGCGGT 2381
QY 1162 TACCGCTCACAGCGGAGCTGGGTTGCTTTCGACGGTGCAGCTCAGCTCGGTGGCGAA 1221
Db 2382 TATCTTCCCTCGCGCGCTTGGCGTGGGATTCGCGATTCACCTCCACGTTTACAGGATTAACCA 2441
QY 1222 ATCACCGCACACTTTGACTCTCATCTGTTGAAATGACCTGCGCGTGTTCGACTTTGAA 1281
Db 2442 ATTCCGCGCTACTACGATTCCTTAATTTGGTAAATGATGCTTTGGGGCCCTGATCGCGCT 2501
QY 1282 ACTGCTGTGCTGTCACAGCGCGCTTGGCTGAGTTTCAACCGTGTCTGCTGTTGCAACC 1341
```

```
Db 2502 ACTGCTATTAAACGCATGAACACGCCCTCAGGGAATGCGCCATCACTGGATTACCTACA 2561
QY 1342 AACATTGGTTT 1352
Db 2562 ACCATTGGTT 2572

RESULT 15
US-08-476-537-1
; Sequence 1, Application US/08476537
; Patent No. 5756290
; GENERAL INFORMATION:
; APPLICANT: Robert Haselkorn and Piotr Gornicki
; TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
; TITLE OF INVENTION: Carboxylase
; NUMBER OF SEQUENCES: 116
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: 321 No. 5756290th Clark Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60610
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,537
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/956,700
; FILING DATE: 10/21/92
; ATTORNEY/AGENT INFORMATION:
; NAME: Thomas E. No. 5756290thrup
; REGISTRATION NUMBER: 33,268
; REFERENCE/DOCKET NUMBER: ARCD:058
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 1-312-744-0090
; TELEFAX: 1-312-755-4489
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3065 base pairs
; TYPE: Nucleic acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: Oligonucleotide
; US-08-476-537-1

Query Match 8.5%; Score 297; DB 1; Length 3065;
Best Local Similarity 54.4%; Pred. No. 1.7e-72;
Matches 691; Conservative 0; Mismatches 565; Indels 15; Gaps 4;

QY 85 TTCAAAAGATCTTGGTAGCAACCGCGGCAATCCGGTCCGCTTTCGCTGCGAGCA 144
Db 1314 TTTGACAAATATTAATTGCCAATCGGGAGAAATAGCGCTGCGCATTTCTCCGCGCTGT 1373
QY 145 CTCGAAACCGGTGAGCCACCGTAGCTATTATACCCCGTGAAGATCGGGGATCAATTCAC 204
Db 1374 GAGGAATGGGATTCGACGATCGCAGTTTCATTCGACTGTTCGACCGGAATGCTCTTTCAT 1433
QY 205 CGCTCTTTTCTTCTGAGCTGTCCGCAATTTGGTACTGAAAGCTCACCAGTCAAGGCGTAC 264
Db 1434 GTCCAACTTCTGACGAAAGCGGTTTGTATTGGC---GAACCTGCTAGCGCTAAAGTTAT 1490
QY 265 CTGACATCGATGAATATTCGGTGCAGCTAAAAAGTTTAAAGCAGATGCTATTATTCACCG 324
Db 1491 TTGAATATCCCAATATATTGCTGGGCTTTAAGCGCAATGCCAGTGTCTTATCATCCT 1550
QY 325 GGATATGGCTTCTGCTGAAATATTCGGTGCAGTGTGCCCGCGAGTGCAGGAAAAACGGCAAT 384
```

Job time : 152 secs

Db 1551 GGGTATGGCTTTTATCTGAAATGCCAAATTTGGGAAATCTGTCTGACCATCATTT 1610
Qy 385 ACTTTTATGGCCCAACCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGTGGGTA 444
Db 1611 GCATTTCAATGGCCCAACCCAGAGGTTATCCGCTCATGGGGACAAATCCACTGCCAAG 1670
Qy 445 ACCGCCGGAAGAGGTTGTCTGCCAGTTTGGCGG--AATCCACCCGAGCAAAAACA 502
Db 1671 GAAACCATGCAAAAAGCTGGTGTACCGACAGTACCGGGTAGTGAAGTTTGGTAGAGACA 1730
Qy 503 TCGATGACATCGTTTAAAGCGCTGAAGGCCA-GACTTACCCCATCTTTGTAAAGGCAGTT 561
Db 1731 GAGCAAGAGGATTAGAACTTGGCGAAGATATTGGCTACCCAGTGATGATCAAGGCCAG 1790
Qy 562 GCCGGTGGTGGCGGACCGGTATGCGCTTTGTTCTTCACTGATGAGCTCCGCAATTTG 621
Db 1791 GCTGGTGGTGGCGGCGGGGTATGCGACTGGTGGATCGGCAGATGAATTTGTCAAACTG 1850
Qy 622 GCAACAGAAGCATCTCGTGAAGCTGAAGCGGCAATCGGCGACGGTTCGGTATATGTCGAA 681
Db 1851 TTCTTAGCCGCCCAAGGTGAAGCTGGTGCAGCCTTTGGTAATGCTGGCGTTTATATAGAA 1910
Qy 682 CGTCTGTGATTAAACCCAGCACATTTGAAGTGCAGATCCTTGGCGATCGCACTGGAGAA 741
Db 1911 AAATTTATTGAACGTCGCGCCCAATTTGAATTTTGGCTGATATTAACGGCAAT 1970
Qy 742 GTTGTACACCTTTATGAACGTTGACTGCTGCTGAGCGTCTGACCAAAAAGTTGTCGAA 801
Db 1971 GTGATTCACTTGGGTGAGAGGATTGCTCAATTGAGCGTCTGTAACCAAAAGTTACTAGAA 2030
Qy 802 ATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGTATCGCATTTTGGGGATGCAGTA 861
Db 2031 GAAGCCCCAGCCAGCCTTGGACTCAGACCTAAGGGAAAAATGGACAAGCGCGGTG 2090
Qy 862 AAGTTCTGCGCTCCATTTGTTACCGGCGCGGACCGTGGAAATCTTGGTCCGATGAA 921
Db 2091 AAAGCGCTCAGTTTATCAATTAGCCGGGCGAGTACTATCGAGTTTGTCTAGATAGA 2150
Qy 922 AAGGGCAACACAGTTTTCATCGAAATGAACCCAGTATCCAGTTGAGCACCGTCACT 981
Db 2151 TCCGGTCAGTTTACTTTATGAGATGACACCCGGATTCAGTAGAAACATCCCGTACT 2210
Qy 982 GAAGAATCACCGAGGTGGACCTGGTGAAGCGGAGATGCGCTTGGCTGCTGTGCAACC 1041
Db 2211 GAGATGGTTACTGGAGTGGATTTATTGGTTGAGCAAAATCAGAAATGCCAAGGGGAAA-- 2268
Qy 1042 TTGAAGGAATGGGTCTGACCCAGATAAGATCAAGACCCACCGTGACGACTGCAAGTGC 1101
Db 2269 -----GACTTAGACTAACTCAAGACCAAGTAGTTTACGCGGTGATGCGATCGAATGT 2321
Qy 1102 CGCATCACCGAAGATCCAAACACCGGCTTCGCGCCAGATACCGGAATATCACCGCG 1161
Db 2322 CGCATCAATCCGAAGACCCAGACACGATTTCCGCGCCAGCACCCGACGCAATTAGCGGT 2381
Qy 1162 TACCGCTCACCGCGGAGCTGGCGTTGCTTTGACGGTGCAGCTCAGCTCGGTGGCGAA 1221
Db 2382 TATCTTCCCTCCGCGCCCTGGCGTGGGATTGACTCCCGAGTTTACCGGATTACCAA 2441
Qy 1222 ATCACCGCACACTTTGATCCATGCTGGTGAATGACCTGCGGTGTTCCGACTTTGAA 1281
Db 2442 ATTCCGCGCTACTACGATTCCTTAATTGGTAAATTGATCGTTTGGGCGCTGATCGCGCT 2501
Qy 1282 ACTGCTGTGCTGTGACAGCGCGGTTGGCTGAGTTGACCGGTGCTGTGTTGCAACC 1341
Db 2502 ACTGCTATTAAACCGCATGAACGCGCCCTCAGGAATGCGGCATCACTGGATTACCTACA 2561
Qy 1342 AACATTGGTTT 1352
Db 2562 ACCATTGGGTT 2572

Search completed: April 5, 2003, 10:03:09